DOCUMENT RESUME

ED 464 893 SO 033 859

AUTHOR	Lapp, Michael S.; Grigg, Wendy S.; Tay-Lim, Brenda S. H.
TITLE	The Nation's Report Card: U.S. History, 2001.
INSTITUTION	National Center for Education Statistics (ED), Washington, DC.
REPORT NO	NCES-2002-483
PUB DATE	2002-05-00
NOTE	179p.; Written in collaboration with Tatyana Petrovicheva and Satwinder Thind. The work upon which this publication is based was performed for the National Center for Education Statistics by the Educational Testing Service.
AVAILABLE FROM	ED Pubs, P.O. Box 1398, Jessup, MD 20794-1398. Tel: 877-433-7827 (Toll Free); Tel: 877-576-7734 (Toll Free; TTY/TDD); Fax: 301-470-1244; e-mail: EdPubs@inet.ed.gov; Web site: http://www.ed.gov/pubs/edpubs.html. For full text: http://nces.ed.gov/nationsreportcard.
PUB TYPE	Numerical/Quantitative Data (110) -- Reports - Research (143)
EDRS PRICE	MF01/PC08 Plus Postage.
DESCRIPTORS	*Academic Achievement; Comparative Analysis; Elementary Secondary Education; *Grade 12; *Grade 4; *Grade 8; Public Schools; School Surveys; *Student Evaluation; *United States History
IDENTIFIERS	*National Assessment of Educational Progress

ABSTRACT

The National Assessment of Educational Progress (NAEP) is the nation's only ongoing representative sample survey of student achievement in core subject areas. In 2001, NAEP conducted a national United States history assessment of fourth-, eighth-, and twelfth-grade students. This report presents a summary of the results of the NAEP 2001 U.S. history assessment for the nation. Results in 2001 are compared to results in 1994, the next most recent year in which the NAEP conducted a U.S. history assessment and the only other assessment year in which the test questions were based on the current framework. Students' performance on the assessment is described in terms of average scores on a 0-500 scale and in terms of the percentage of students attaining 3 achievement levels: (1) Basic; (2) Proficient; and (3) Advanced. The report provides results for subgroups of students at the three grade levels defined by various background and contextual characteristics. Appended are an overview of the procedures used in the NAEP assessment and data collection. (BT)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 464 893

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

National Center for Education Statistics

The Nation's Report Card
U.S. History 2001

SO 033 859

BEST COPY AVAILABLE

NAEP
U.S. Department of Education
Office of Educational Research and Improvement
2
NCES 2002-483

What is The Nation's Report Card?

THE NATION'S REPORT CARD, the National Assessment of Educational Progress (NAEP), is the only nationally representative and continuing assessment of what America's students know and can do in various subject areas. Since 1969, assessments have been conducted periodically in reading, mathematics, science, writing, history, geography, and other fields. By making objective information on student performance available to policymakers at the national, state, and local levels, NAEP is an integral part of our nation's evaluation of the condition and progress of education. Only information related to academic achievement is collected under this program. NAEP guarantees the privacy of individual students and their families.

NAEP is a congressionally mandated project of the National Center for Education Statistics, the U.S. Department of Education. The Commissioner of Education Statistics is responsible, by law, for carrying out the NAEP project through competitive awards to qualified organizations. NAEP reports directly to the Commissioner, who is also responsible for providing continuing reviews, including validation studies and solicitation of public comment, on NAEP's conduct and usefulness.

In 1988, Congress established the National Assessment Governing Board (NAGB) to formulate policy guidelines for NAEP. The Board is responsible for selecting the subject areas to be assessed from among those included in the National Education Goals; for setting appropriate student performance levels; for developing assessment objectives and test specifications through a national consensus approach; for designing the assessment methodology; for developing guidelines for reporting and disseminating NAEP results; for developing standards and procedures for interstate, regional, and national comparisons; for determining the appropriateness of test items and ensuring they are free from bias; and for taking actions to improve the form and use of the National Assessment.

The National Assessment Governing Board

Mark D. Musick, Chair
President
Southern Regional Education Board
Atlanta, Georgia

Michael T. Nettles, Vice Chair
Professor of Education
University of Michigan
Ann Arbor, Michigan

Melanie A. Campbell
Fourth-Grade Teacher
Topeka, Kansas

Honorable Wilmer S. Cody
Former Commissioner of Education
State of Kentucky
Frankfort, Kentucky

Daniel A. Domenech
Superintendent of Schools
Fairfax County Public Schools
Fairfax, Virginia

Edward Donley
Former Chairman
Air Products & Chemicals, Inc.
Allentown, Pennsylvania

Thomas H. Fisher
Director
Student Assessment Services
Florida Department of Education
Tallahassee, Florida

Edward H. Haertel
Professor, School of Education
Stanford University
Stanford, California

Juanita Haugen
Local School Board Member
Pleasanton, California

Honorable Dirk Kempthorne
Governor of Idaho
Boise, Idaho

Honorable Nancy Kopp
State Legislator
Annapolis, Maryland

Honorable Ronnie Musgrove
Governor of Mississippi
Jackson, Mississippi

Roy M. Nageak, Sr.
First Vice-Chair
Alaska Board of Education and Early Development
Barrow, Alaska

Debra Paulson
Eighth-Grade Mathematics Teacher
El Paso, Texas

Honorable Jo Ann Pottorff
State Legislator
Wichita, Kansas

Diane Ravitch
Research Professor
New York University
New York, New York

Sister Lourdes Sheehan, R.S.M.
Secretary for Education
United States Catholic Conference
Washington, DC

John H. Stevens
Executive Director
Texas Business and Education Coalition
Austin, Texas

Migdania D. Vega
Principal
Coral Way Elementary Bilingual School
Miami, Florida

Deborah Voltz
Assistant Professor
Department of Special Education
University of Louisville
Louisville, Kentucky

Honorable Michael E. Ward
State Superintendent of Public Instruction
North Carolina Public Schools
Raleigh, North Carolina

Marilyn A. Whirry
Twelfth-Grade English Teacher
Manhattan Beach, California

Dennie Palmer Wolf
Director, Annenberg Institute
Brown University
Providence, Rhode Island

Grover J. Whitehurst (Ex-Officio)
Assistant Secretary of Education
Office of Educational Research and Improvement
U.S. Department of Education
Washington, DC

Roy Truby
Executive Director, NAGB
Washington, DC

National Center for Education Statistics

The Nation's Report Card
U.S. History 2001

Michael S. Lapp

Wendy S. Grigg

Brenda S.-H. Tay-Lim

in collaboration with

Tatyana Petrovicheva

Satwinder Thind

May 2002

U.S. Department of Education
Office of Educational Research and Im[illegible]ent

NCES 2002–483

U.S. Department of Education
Rod Paige
Secretary

Office of Educational Research and Improvement
Grover J. Whitehurst
Assistant Secretary

National Center for Education Statistics
Gary W. Phillips
Deputy Commissioner

May 2002

SUGGESTED CITATION
U.S. Department of Education. Office of Educational Research and Improvement. National Center for Education Statistics. *The Nation's Report Card: U.S. History 2001*, NCES 2002–483, by M. S. Lapp, W. S. Grigg, & Brenda S.-H. Tay-Lim. Washington, DC: 2002.

FOR MORE INFORMATION
Content contact:
Janis Brown
202–502–7419

To obtain single copies of this report, limited number of copies available, or ordering information on other U.S. Department of Education products, call toll free 1–877–4ED-PUBS (877–433–7827), or write:

Education Publications Center (ED Pubs)
U.S. Department of Education
P.O. Box 1398
Jessup, MD 20794–1398

TTY/TDD 1–877–576–7734
FAX 301–470–1244

Online ordering via the Internet: *http://www.ed.gov/pubs/edpubs.html*
Copies also are available in alternate formats upon request.
This report also is available on the World Wide Web: *http://nces.ed.gov/nationsreportcard*

The work upon which this publication is based was performed for the National Center for Education Statistics by Educational Testing Service.

Table of Contents

Executive Summary ix

Chapter 1

NAEP 2001 U.S. History Assessment 1

Introduction 1
Overview of the 2001 National Assessment of Educational Progress 2
The U.S. History Framework 2
U.S. History Assessment Instruments 5
Description of School and Student Samples 6
Reporting the Assessment Results 7
The Setting of Achievement Levels 8
Achievement-Level Descriptions for Each Grade 9
The Trial Status of Achievement Levels 13
Interpreting NAEP Results 15
Overview of the Remaining Report 16

Chapter 2

Average Scale Scores and Achievement-Level Results 17

Overview 17
Average Scale Score Results 18
Scale Scores by Percentile 19
Achievement-Level Results 20

Chapter 3

Average Scale Scores and Achievement-Level Results for Selected Subgroups 23

Gender 24
Race/Ethnicity 27
Region of the Country 32
Parents' Highest Level of Education 36
Type of School 40
Type of Location 45
Free/Reduced-Price School Lunch Program Eligibility 47

Chapter 4

Becoming a More Inclusive National Assessment 49

Two Sets of 2001 NAEP U.S. History Results 50

Results for the Nation 53

National Results by Gender 55

National Results by Race/Ethnicity 57

Chapter 5

Sample Assessment Questions and Student Responses 59

Grade 4 Sample Assessment Question Results 60

Grade 8 Sample Assessment Question Results 68

Grade 12 Sample Assessment Question Results 75

Maps of Selected Item Descriptions on the NAEP U.S. History Scale—Grades 4, 8, and 12 84

Chapter 6

Classroom Contexts for Learning 89

Time Spent on Social Studies 89

State and Local Standards 91

Instructional Activities 92

Use of Technology 99

Appendix A

Overview of Procedures Used for the NAEP 2001 U.S. History Assessment 103

Appendix B

Data Appendix 133

Appendix C

Members of the NAEP U.S. History Standing Committee 165

Acknowledgments 166

Chapter 1: Figures and Tables

Figure 1.1: Descriptions of the four U.S. history themes 3

Table 1.1: Distribution of assessment time across historical themes, grades 4, 8, and 12: 2001 4

Table 1.2: Distribution of assessment time across historical periods, grades 4, 8, and 12: 2001 5

Figure 1.2: Policy definitions of the three NAEP achievement levels 8

Figure 1.3: Descriptions of NAEP U.S. history achievement levels for grade 4 10

Figure 1.4: Descriptions of NAEP U.S. history achievement levels for grade 8 11

Figure 1.5: Descriptions of NAEP U.S. history achievement levels for grade 12 12

Chapter 2: Figures and Tables

Figure 2.1: Average U.S. history scale scores, grades 4, 8, and 12: 1994 and 2001 18

Figure 2.2: U.S. history scale score percentiles, grades 4, 8, and 12: 1994 and 2001 19

Figure 2.3: Percentage of students within and at or above U.S. history achievement levels, grades 4, 8, and 12: 1994 and 2001 21

Chapter 3: Figures and Tables

Figure 3.1: Average U.S. history scale scores by gender, grades 4, 8, and 12: 1994 and 2001 24

Figure 3.2: Differences in average U.S. history scale scores by gender, grades 4, 8, and 12: 1994 and 2001 25

Figure 3.3: Percentage of students within and at or above U.S. history achievement levels by gender, grades 4, 8, and 12: 1994 and 2001 ... 26

Figure 3.4: Average U.S. history scale scores by race/ethnicity, grades 4, 8, and 12: 1994 and 2001 27

Figure 3.5: Differences in average U.S. history scale scores by race/ethnicity, grades 4, 8, and 12: 1994 and 2001 28

Figure 3.6a: Percentage of students within and at or above U.S. history achievement levels by race/ethnicity, grade 4: 1994 and 2001 29

Figure 3.6b: Percentage of students within and at or above U.S. history achievement levels by race/ethnicity, grade 8: 1994 and 2001 30

Figure 3.6c: Percentage of students within and at or above U.S. history achievement levels by race/ethnicity, grade 12: 1994 and 2001 31

Figure 3.7: Average U.S. history scale scores by region of the country, grades 4, 8, and 12: 1994 and 2001 32

Figure 3.8a: Percentage of students within and at or above U.S. history achievement levels by region of the country, grade 4: 1994 and 2001 33

Figure 3.8b: Percentage of students within and at or above U.S. history achievement levels by region of the country, grade 8: 1994 and 2001 34

Figure 3.8c: Percentage of students within and at or above U.S. history achievement levels by region of the country, grade 12: 1994 and 2001 35

Figure 3.9: Average U.S. history scale scores by parents' highest level of education, grades 8 and 12: 1994 and 2001 37

Figure 3.10a: Percentage of students within and at or above U.S. history achievement levels by parents' highest level of education, grade 8: 1994 and 2001 38

Figure 3.10b: Percentage of students within and at or above U.S. history achievement levels by parents' highest level of education, grade 12: 1994 and 2001 ... 39

Figure 3.11: Average U.S. history scale scores by type of school, grades 4, 8, and 12: 1994 and 2001 40

Figure 3.12a: Percentage of students within and at or above U.S. history achievement levels by type of school, grade 4: 1994 and 2001 42

Figure 3.12b: Percentage of students within and at or above U.S. history achievement levels by type of school, grade 8: 1994 and 2001 43

Figure 3.12c: Percentage of students within and at or above U.S. history achievement levels by type of school, grade 12: 1994 and 2001 44

Table 3.1: Average U.S. history scale scores by type of location, grades 4, 8, and 12: 2001 45
Figure 3.13: Percentage of students within and at or above U.S. history achievement levels by type of school location, grades 4, 8, and 12: 2001 46
Table 3.2: Average U.S. history scale scores by student eligibility for Free/Reduced-Price School Lunch program, grades 4, 8, and 12: 2001 47
Figure 3.14: Percentage of students within and at or above U.S. history achievement levels by Free/Reduced-Price School Lunch program eligibility, grades 4, 8, and 12: 2001 48

Chapter 4: Figures and Tables

Figure 4.1: The two sets of NAEP results based on a split-sample design 52
Table 4.1: National average U.S. history scale scores by type of results, grades 4, 8, and 12: 2001 54
Table 4.2: Percentage of students within and at or above U.S. history achievement levels by type of results, grades 4, 8, and 12: 2001 55
Table 4.3: National average U.S. history scale scores by gender and type of results, grades 4, 8, and 12: 2001 55
Table 4.4: Percentage of students within and at or above U.S. history achievement levels by gender and type of results, grades 4, 8, and 12: 2001 56
Table 4.5: National average U.S. history scale scores by race/ethnicity and type of results, grades 4, 8, and 12: 2001 57
Table 4.6: Percentage of students within and at or above U.S. history achievement levels by race/ethnicity and type of results, grades 4, 8, and 12: 2001 58

Chapter 5: Figures and Tables

Table 5.1: Grade 4 Sample Question 1 Results (Multiple-Choice) 60
Table 5.2: Grade 4 Sample Question 2 Results (Multiple-Choice) 61
Table 5.3: Grade 4 Sample Question 3 Results (Multiple-Choice) 62
Table 5.4: Grade 4 Sample Question 4 Results (Extended Constructed-Response) ... 64
Table 5.5: Grade 8 Sample Question 5 Results (Multiple-Choice) 68
Table 5.6: Grade 8 Sample Question 6 Results (Multiple-Choice) 69
Table 5.7: Grade 8 Sample Question 7 Results (Short Constructed-Response) 70
Table 5.8: Grade 8 Sample Question 8 Results (Short Constructed-Response) 73
Table 5.9: Grade 12 Sample Question 9 Results (Multiple-Choice) 75
Table 5.10: Grade 12 Sample Question 10 Results (Multiple-Choice) 76
Table 5.11: Grade 12 Sample Question 11 Results (Extended Constructed-Response) .. 78
Table 5.12: Grade 12 Sample Question 12 Results (Extended Constructed-Response) .. 82
Figure 5.1: Grade 4 Item Map
Map of selected item descriptions on the National Assessment of Educational Progress U.S. history scale for grade 4 85
Figure 5.2: Grade 8 Item Map
Map of selected item descriptions on the National Assessment of Educational Progress U.S. history scale for grade 8 86
Figure 5.3: Grade 12 Item Map
Map of selected item descriptions on the National Assessment of Educational Progress U.S. history scale for grade 12 87

Chapter 6: Figures and Tables

Table 6.1: Percentage of students and average U.S. history scale scores by teachers' reports on the amount of time spent on social studies in a typical week at grade 4: 2001 90

Table 6.2: Percentage of students and average U.S. history scale scores by teachers' reports on the use of state/local standards in planning instruction at grades 4 and 8: 2001 92

Table 6.3: Percentage of students and average U.S. history scale scores by teachers' reports on frequency of classroom activities at grades 4 and 8: 1994 and 2001 94

Table 6.4: Percentage of students and average U.S. history scale scores by students' reports on frequency of classroom activities at grade 12: 1994 and 2001 97

Table 6.5: Percentage of students and average U.S. history scale scores by students' reports on computer use at grade 4: 2001 100

Table 6.6: Percentage of students and average U.S. history scale scores by students' reports on computer use at grades 8 and 12: 2001 101

Executive Summary

The National Assessment of Educational Progress (NAEP) is the nation's only ongoing representative sample survey of student achievement in core subject areas. In 2001, NAEP conducted a national U.S. history assessment of fourth-, eighth-, and twelfth-grade students.

Authorized by Congress and administered by the National Center for Education Statistics (NCES) in the U.S. Department of Education, NAEP regularly reports to the public on the educational progress of students in grades 4, 8, and 12. This report presents the results of the NAEP 2001 U.S. history assessment for the nation. Results in 2001 are compared to results in 1994, the next most recent year in which NAEP conducted a U.S. history assessment and the only other assessment year in which the test questions were based on the current framework. Students' performance on the assessment is described in terms of average scores on a 0–500 scale and in terms of the percentage of students attaining three achievement levels: *Basic*, *Proficient*, and *Advanced*. The achievement levels are performance standards adopted by the National Assessment Governing Board (NAGB) as part of its statutory responsibilities. They are collective judgments of what students should know and be able to do.

The Nation's Report Card

Major Findings at Grades 4, 8, and 12

Results for Student Subgroups

Becoming a More Inclusive NAEP

Classroom Contexts for Learning

As provided by law, the Deputy Commissioner of Education Statistics, upon review of a congressionally mandated evaluation of NAEP, has determined that the achievement levels are to be used on a trial basis and should be interpreted and used with caution. However, both the Deputy Commissioner and NAGB believe these performance standards are useful for understanding trends in student achievement. They have been widely used by national and state officials as a common yardstick of academic performance.

In addition to providing average scores and achievement-level performance in U.S. history for the nation's fourth-, eighth-, and twelfth-graders, this report provides results for subgroups of students at those grade levels defined by various background and contextual characteristics.

A summary of major findings from the NAEP 2001 U.S. history assessment is presented on the following pages. In interpreting NAEP results, it should be noted that every test score has a standard error—a range of a few points plus or minus the score—that includes components of sampling error and measurement error. Statistical tests that factor in these standard errors are used to determine whether the differences between average scores are significant. Only statistically significant differences are cited in this report. Readers are also cautioned against making causal inferences based on NAEP results. Differences in performance between subgroups of students, for example, reflect a variety of socioeconomic and educational factors.

Major Findings at Grades 4, 8, and 12

- Average U.S. history scores for fourth- and eighth-graders were higher in 2001 than in 1994, while the performance of twelfth-graders remained relatively stable.
- Score increases were evident among the lower-performing students at grade 4 (at the 10th and 25th percentiles) and for both lower- and higher-performing students at grade 8 (25th, 75th, and 90th percentiles).
- Results of the 2001 U.S. history assessment show 18 percent of fourth-graders, 17 percent of eighth-graders, and 11 percent of twelfth-graders performing at or above the *Proficient* level—identified by NAGB as the level at which all students should perform.
- At grade 4, the percentage of students performing at or above *Basic* in 2001 was higher than in 1994. At grade 8, the percentages of students performing at or above *Basic*, at or above *Proficient*, and at *Advanced* increased between 1994 and 2001. At grade 12, however, the percentages performing at or above each level remained the same as in 1994.

Results for Student Subgroups

In addition to overall results, NAEP reports on the performance of various subgroups of students. Observed differences between student subgroups in NAEP U.S. history performance reflect a range of socioeconomic and educational factors not addressed in this report or by NAEP.

Gender

- Any apparent differences in the average scores of male and female students in 2001 were not statistically significant at any of the three grades.
- At grade 4, both male and female students had higher average scores in 2001 than in 1994. At grade 8, the average score of males increased between 1994 and 2001, while the performance of females remained stable.

Race/Ethnicity

- In 2001, the average scores of White students were higher than those of Black, Hispanic, and American Indian students at all three grades. Asian/Pacific Islander students scored higher than Black and Hispanic students across the grades as well.
- At grade 4, both White students and Black students had higher average scores in 2001 than in 1994. At grade 8, only White students showed a gain since 1994. At grade 12, only Hispanic students had higher average scores in 2001 than in 1994.
- The 2001 results show a narrowing of the score point difference between White students and Black students at grade 4, and between White students and Hispanic students at grade 12.

Region of the Country

- Fourth- and eighth-grade students in the Northeast, Southeast, and Central regions all had higher average scores than students in the West. Fourth- and eighth-grade students in the Central region outperformed their peers in the Southeast. There was no statistically significant difference in the performance of twelfth-graders from various regions of the country.
- At grade 4, only the Northeast region showed a gain in the U.S. history average score since 1994. At grade 8, the only increase occurred in the Southeast region.

Parents' Highest Level of Education

- The 2001 results show a clear positive relationship overall between parental education level and the performance of eighth- and twelfth-graders.
- At grade 8, the average score of students whose parents graduated from college was higher in 2001 than in 1994. At grade 12, there was an increase in the average score of students whose parents did not finish high school.

Type of School

- The 2001 results show that public school students at all three grades had lower average U.S. history scores than their peers attending nonpublic schools.
- Average scores among both fourth- and eighth-grade public school students were higher in 2001 than in 1994.

Type of Location

- At grades 4 and 8, students attending schools in rural and urban fringe locations had higher average scores than students in central city schools. At grade 12, students attending schools in urban fringe locations had higher scores than students in both rural and central city locations.

Eligibility for Free/Reduced-Price School Lunch Program

- At every grade, the average score of students who were eligible for the Free/Reduced-Price School Lunch program was lower than the average score of students who were not eligible for the program (i.e., those not meeting the poverty guidelines).

Becoming a More Inclusive NAEP

In the 2001 U.S. history assessment, the NAEP program used a split-sample design, so that trends in students' history achievement could be reported across assessment years and, at the same time, the program could continue to examine the effects of including special-needs students assessed with accommodations. While most of the results in this report include only the performance of students assessed without accommodations, the report also presents an overview of a second set of results that include the performance of special-needs students who required and were provided accommodations during the assessment administration.

- At grade 8, the average score when accommodations were permitted was lower than the average score when accommodations were not permitted. However, there were no statistically significant differences between average scores in the accommodations-permitted results and the accommodations-not-permitted results at grades 4 and 12.

Classroom Contexts for Learning

NAEP collects information about the contexts for student learning by administering questionnaires to assessed students, their teachers, and their school administrators. Using the student as the unit of analysis, NAEP examines the relationship between selected contextual variables drawn from these questionnaires and students' average scores on the U.S. history assessment.

Time Spent on Social Studies

- In 2001, fourth-graders whose teachers reported spending more than 180 minutes on social studies instruction in a typical week had higher average scores than those whose teachers reported spending less time.

State and Local Standards

- About two-thirds of the fourth- and eighth-graders assessed had teachers who reported that they used state or local standards to a large extent in planning social studies instruction. There were no statistically significant differences in students' performance at either grade 4 or grade 8 based on the extent to which teachers reported using such standards in planning instruction.

Instructional Activities

- A large majority of fourth-graders had teachers who reported having them read material from a textbook on a daily or weekly basis. Reading from a textbook daily was associated with higher average scores than was doing so on a weekly or monthly basis.
- Eighth-graders whose teachers reported using primary historical documents such as letters, diaries, or essays written by historical figures, on a weekly basis had higher average scores than those whose teachers did so less frequently.
- Twelfth-graders who reported never reading extra material, such as biographies or historical stories, scored lower, on average, than those who reported doing so a few times a year or more often.

Use of Technology

- A strong positive association was evident between using computers for conducting research and for writing reports and performance at grades 8 and 12.
- Students in grades 4, 8, and 12 who reported daily general use of computers at school for social studies or history had lower average scores than those who reported less frequent general use. It should be noted that relatively few students reported using a computer for history or social studies.

1 NAEP 2001 U.S. History Assessment

Introduction

Knowledge of United States history is an important component of effective citizenship. Having a thorough grasp of our country's struggles and achievements better enables young people to make informed and intelligent decisions about contemporary issues. Nourishing the curiosity children exhibit about major events, customs and institutions, and the families and individuals that comprise United States history, creates a valuable resource for our nation's future.

Chapter Focus

What is the NAEP U.S. history assessment?

How does the NAEP U.S. history assessment measure and report student progress?

Efforts to improve the rigor and quality of history education have been an important element of the standards-based educational reform movement of recent years. This emphasis on high expectations has been demonstrated by the rapid increase in the number of states that have, within the last 10 years, established content standards for history or social studies; between 1995 and 2000 the number of states with such standards increased from 20 to 46.[1] The renewed interest in history instruction underscores the need for accurate information about what students know and can do in U.S. history.

Chapter Contents

Overview

U.S. History Framework

U.S. History Assessment

School and Student Samples

Reporting Results

NAEP Achievement Levels

Interpreting NAEP Results

This Report

[1] Council of Chief State School Officers. (2000). *Key state education policies on K-12 education: 2000.* (table 13, p. 23). Washington, DC: Author.

BEST COPY AVAILABLE

Overview of the 2001 National Assessment of Educational Progress

For over 30 years, the National Assessment of Educational Progress (NAEP) has been authorized by Congress to collect, analyze, and report reliable and valid information about what American students know and can do in core subject areas. NAEP assesses the performance of public and nonpublic school students in grades 4, 8, and 12. In 2001, student performance in U.S. history and geography was assessed at all three grades. This report deals only with the results of the U.S. history assessment.

All NAEP assessments are based on content frameworks developed through a national consensus process. The NAEP 2001 U.S. history assessment was the second administration of an assessment based on the *NAEP U.S. History Framework*, which was originally developed for the 1994 assessment.[2] In both 1994 and 2001, assessments based on the framework were administered to national samples of fourth-, eighth-, and twelfth-graders.

This report describes the results of the 2001 U.S. history assessment at grades 4, 8, and 12 and compares results in 2001 to those in 1994. Comparisons across assessment years are possible because the assessments were developed under the same basic framework and share a common set of U.S. history questions. In addition, the populations of students were sampled and assessed using comparable procedures.

The U.S. History Framework

Although U.S. history was assessed by NAEP in 1986 and 1988, a rigorous new NAEP U.S. History Framework was developed for the 1994 assessment. The new framework provided the operational specifications for both the 1994 and 2001 assessments. The development of the framework was managed by the Council of Chief State School Officers (CCSSO) under the direction of the National Assessment Governing Board (NAGB). Approximately 50 professional historians, educators, administrators, and other interested individuals worked to achieve consensus on the general goals as well as the specific language of the framework. In addition, several hundred educational experts and interested members of the public contributed to the process, either by participating in public hearings or by reviewing drafts. The framework document produced by this consensus process called for the assessment of a broad range of outcomes. It represented an ambitious vision both of what students should know and be able to do in U.S. history, and of the ways in which those competencies should be tested.

---

2 National Assessment Governing Board. (1993). *U.S. history framework for the 1994 National Assessment of Educational Progress*. Washington, DC: Author.

The framework is organized around three concepts or dimensions: major themes of U.S. history, chronological periods, and ways of knowing and thinking about U.S. history with the four themes providing the core organizing structure of the framework. The themes were intended to ensure that all major branches of historical study were covered and that emphasis on various areas was balanced. The themes are also used to define the subscales that make up the NAEP U.S. history composite scale. (See appendix A for more information on how the scale was constructed.) Figure 1.1 provides descriptions of each theme.

Figure 1.1 Descriptions of the four U.S. history themes

Theme	Descriptions
Change and Continuity in American Democracy: Ideas, Institutions, Practices, and Controversies	This theme concerns the development of American political democracy from colonial times to the present. It covers political events that shaped American democracy, such as the American Revolution, the Civil War, the fight for civil rights, as well as the core ideas and principles that underlie our institutions. This theme covers students' knowledge of the founding of the nation, the writing of the Constitution, and other fundamental components of the nation's political history. At the same time, it calls for evaluating students' understanding of the role that major political ideas and conflicts have played at different points in our history.
The Gathering and Interactions of Peoples, Cultures, and Ideas	This theme is broadly defined because it covers a vast component of U.S. history: the interactions among the peoples and cultures of many countries, racial and ethnic groups, and religious traditions that have contributed to the development of American society. This theme covers immigration, cultural developments, patterns of social organization, and changing roles of men and women.
Economic and Technological Changes and Their Relation to Society, Ideas, and the Environment	This theme focuses on the economic history of the nation and its development from a rural, agricultural society to an urban, industrialized superpower. It also covers the roles of geography and of developments in science and technology in bringing about socio-economic change.
The Changing Role of America in the World	This theme calls for coverage of the many factors—political ideas, economic interests, public opinion—that have shaped American foreign policy. It also addresses specific interactions between the United States and other nations and domestic consequences of developments in foreign policy.

SOURCE: National Assessment Governing Board. U.S. History Framework for the 1994 and 2001 National Assessment of Educational Progress.

BEST COPY AVAILABLE

Eight periods provide chronological structure for the many issues included in the four themes. These periods focus attention on several major eras of U.S. history. They overlap at some points because they were conceived to ensure thorough coverage of major trends and events. The historical periods are not used as subscales, but rather were used in the assessment development process to ensure appropriate chronological coverage. The periods are as follows:

- Three Worlds and Their Meeting in the Americas (Beginnings to 1607)
- Colonization, Settlement, and Communities (1607 to 1763)
- The Revolution and the New Nation (1763 to 1815)
- Expansion and Reform (1801 to 1861)
- Crisis of the Union: Civil War and Reconstruction (1850 to 1877)
- The Development of Modern America (1865 to 1920)
- Modern America and the World Wars (1914 to 1945)
- Contemporary America (1945 to Present)

The percentages of assessment time allotted to each theme and period described in the framework are presented in tables 1.1 and 1.2, respectively. It should be noted that these percentages vary somewhat from the targeted distribution. (See appendix A, Table A.1, for a comparison of the actual and targeted distributions.)

Table 1.1 Assessment Time by Historical Themes

Distribution of assessment time across historical themes, grades 4, 8, and 12: 2001

	Change and Continuity in American Democracy: Ideas, Institutions, Practices, and Controversies	*The Gathering and Interactions of Peoples, Cultures, and Ideas*	*Economic and Technological Changes and Their Relation to Society, Ideas, and the Environment*	*The Changing Role of America in the World*
Grade 4	25%	32%	32%	12%
Grade 8	30%	32%	25%	13%
Grade 12	28%	26%	22%	25%

NOTE: Percentages may not add to 100 due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

BEST COPY AVAILABLE

Table 1.2 Assessment Time by Historical Periods

Distribution of assessment time across historical periods, grades 4, 8, and 12: 2001

	Beginnings to 1607	1607 to 1763	1763 to 1815	1801 to 1861	1850 to 1877	1865 to 1920	1914 to 1945	1945 to Present	No period*
Grade 4	13%	14%	11%	14%	9%	13%	7%	10%	9%
Grade 8	3%	7%	18%	9%	8%	18%	14%	10%	12%
Grade 12	1%	8%	9%	11%	7%	11%	32%†	20%	1%

NOTE: Percentages may not add to 100 due to rounding.
* A number of questions (especially at grades 4 and 8) assessed general historical skills and could not be appropriately classified in any given period.
† This high percentage is largely a function of the 50-minute theme block, which focuses on the World War II homefront. If this block were excluded from calculations, questions covering this period would make up 22% of the assessment at grade 12.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

The framework also considers ways of knowing and thinking about U.S. history. These are divided into two general cognitive domains that were used as a guide in exercise development. The two domains and their definitions are as follows:

☐ Historical Knowledge and Perspective

This domain includes knowing and understanding people, events, concepts, themes, movements, contexts, and historical sources; sequencing events; recognizing multiple perspectives and seeing an era or movement through the eyes of different groups; and developing a general conceptualization of U.S. history.

☐ Historical Analysis and Interpretation

This domain includes explaining issues, identifying historical patterns; establishing cause-and-effect relationships; finding value statements; establishing significance; applying historical knowledge; weighing evidence to draw sound conclusions; making defensible generalizations; and rendering insightful accounts of the past.

U.S. History Assessment Instruments

As the only federally authorized ongoing assessment of U.S. history achievement, the NAEP assessment must reflect the spirit of the framework as well as the specifications provided by it. In order to achieve those goals, the assessment development process involves stages of review by teachers and teacher educators, state officials, and measurement experts. All components of the assessment are evaluated for curricular relevance, developmental appropriateness, and fairness. Final approval of NAEP test questions is given by the National Assessment Governing Board. A list of the U.S. history development committee members for the 2001 assessment is provided in appendix C.

BEST COPY AVAILABLE

The 2001 U.S. history assessment booklets at grades 4, 8, and 12 generally contained three or four sections: a set of general background questions, a set of subject-specific background questions dealing largely with the student's use of technology, and one or two sets, or "blocks," of cognitive questions assessing knowledge and skills in U.S. history. At grades 8 and 12, students were given either two 25-minute blocks or one 50-minute block. At grade 4, however, only 25-minute blocks were used.

At grade 4 a total of six sections, or "blocks", of cognitive questions were given, while at grades 8 and 12 nine blocks were administered. In addition to the cognitive questions, each assessment booklet also included a set of background questions that asked students to give information about their school practices, such as the frequency with which they use computers at school for social studies, do research projects using a CD or the Internet, and write reports. The assessment time for each grade was 50 minutes plus the 10–15 minutes needed to complete the background questions.

Each block of cognitive questions consisted of both multiple-choice and "constructed-response" questions. ("Constructed-response" is the term used to describe test questions in which students write a response, as distinct from multiple-choice questions, in which students choose an answer from one of several options.) Typically, a block will contain about 16–18 questions, but there is considerable variation depending on the balance between multiple-choice and constructed-response questions. Overall, more than 50 percent of student assessment time was devoted to the latter question type. Two types of constructed-response questions were used:

- short constructed-response questions that required students to provide answers, usually in response to a text or visual stimulus, in one or two sentences; and
- extended constructed-response questions that required students to provide answers of a paragraph or more in length

The 50-minute blocks administered at grades 8 and 12 included questions focusing on a particular theme, and included extended constructed-response questions requiring students to synthesize elements from various primary sources. The total number of test questions used in grades 4, 8, and 12 were 94, 145, and 154, respectively. Each student answered only a small portion of the total number of questions. Additional information about the design of the 2001 U.S. history assessment is presented in appendix A.

Description of School and Student Samples

The NAEP 2001 U.S. history assessment included representative samples of both public and nonpublic schools. Approximately 7,000 fourth-graders, 11,000 eighth-graders, and 11,000 twelfth-graders were assessed. The number of schools in the reporting sample were 365 at grade four, 369 at grade eight, and 374 at grade twelve. Each selected school that participated in the assessment and each student assessed represent a portion of the population of interest. For additional information on sample sizes and participation rates, see appendix A.

This report contains two different sets of national results based on two reporting samples that differed in terms of whether or not accommodations were made available to special-needs students. The national results presented in chapters 2, 3, 5, and 6 of this report are based on a nationally representative sample that included special-needs students only if they could be assessed meaningfully without accommodations. These results can be compared to those from 1994, because accommodations were also not made available in that assessment year. Chapter 4 presents a second set of national results from 2001 for a representative sample that includes the performance of students who required and were provided with accommodations (e.g., bilingual dictionary, extended time, small group testing). No comparison of these results to those from 1994 can be made because of the inclusion of these accommodated special-needs students.

In the sample that did not permit accommodations, 7 percent of fourth-graders, 8 percent of eighth-graders, and 4 percent of twelfth-graders were excluded from the U.S. history assessment in 2001. School staff familiar with these students made the determination that these students could not be assessed meaningfully without accommodations, because of their disability and/or limited English proficiency. In 1994, 5 percent at both the fourth and eighth grades, and 3 percent at the twelfth grade were excluded. Additional information regarding exclusion rates is also provided in appendix A.

Reporting the Assessment Results

Student performance on the NAEP U.S. history assessment is presented in two ways: as average scores on the NAEP U.S. history scale, and in terms of the percentage of students attaining NAEP U.S. history achievement levels. The average scale scores are a measure of students' performance on the assessment. The achievement level results indicate the degree to which student performance meets expectations of what they should know and be able to do.

Average scale score results are presented on the NAEP U.S. history composite scale, which ranges from 0-500. Students' responses on the NAEP 2001 U.S. history assessment were analyzed to determine the percentages of students that responded correctly to each multiple-choice question and the percentages of students that responded at each score level for the constructed-response questions. Scales that summarize results for each of the four themes described earlier were created. The composite scale is a weighted average of the separate subscales for the four themes. The weight for each theme corresponds to the theme's relative importance in the NAEP U.S. history framework. A full description of NAEP scale procedures can be found in the forthcoming *NAEP 2001 Technical Report*.

Achievement-level results are presented in terms of U.S. history achievement levels as authorized by the NAEP legislation and adopted by the National Assessment Governing Board.[3] For each grade tested,

[3] No Child Left Behind Act of 2001: Reauthorization of the Elementary and Secondary Education Act. Pub. L. No. 107-110 (H.R. 1).

National Assessment of Educational Progress Improvement Act of 1988. Pub. L. No. 100-297, 20, U.S.C. 1211.

BEST COPY AVAILABLE

NAGB has adopted three achievement levels: *Basic*, *Proficient*, and *Advanced*. For reporting purposes, the achievement-level cut scores are placed on the U.S. history scale, resulting in four ranges: below *Basic*, *Basic*, *Proficient*, and *Advanced*.

The Setting of Achievement Levels

The 1988 NAEP legislation that created the National Assessment Governing Board directed the Board to identify "appropriate achievement goals...for each subject area" that NAEP measures.[4] The 2001 NAEP reauthorization reaffirmed many of the Board's statutory responsibilities, including developing "appropriate student achievement levels for each grade or age in each subject area to be tested . . ."[5] In order to follow this directive and achieve the mandate of the 1988 statute to "improve the form and use of NAEP results," NAGB undertook the development of student performance standards called "achievement levels." Since 1990 the Board has adopted achievement levels in mathematics, reading, U.S. history, geography, science, writing, and civics.

The Board defined three levels for each grade: *Basic*, *Proficient*, and *Advanced*. The *Basic* level denotes partial mastery of the knowledge and skills that are fundamental for proficient work at a given grade. The *Proficient* level represents solid academic performance. Students reaching this level demonstrate competency over challenging subject matter. The *Advanced* level presumes mastery of both the *Basic* and *Proficient* levels. Figure 1.2 presents the policy definitions of the achievement levels that apply across all grades and subject areas. The policy definitions guided the development of the U.S. history achievement levels, as well as the achievement levels established in all other subject areas. Adopting three levels of achievement for each grade signals the importance of looking at more than one standard of performance. The Board believes, however, that all students should reach the *Proficient* level: the *Basic* level is not the desired goal, but rather represents partial mastery that is a step toward *Proficient*.

Figure 1.2 Policy definitions of the three NAEP achievement levels

Achievement Levels	
Basic	This level denotes partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade.
Proficient	This level represents solid academic performance for each grade assessed. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.
Advanced	This level signifies superior performance.

SOURCE: National Assessment Governing Board. U.S. History Framework for the 1994 and 2001 National Assessment of Educational Progress.

4 National Assessment of Educational Progress Improvement Act of 1988. Pub. L. No. 100-297, 20, U.S.C. 1211.

5 No Child Left Behind Act of 2001: Reauthorization of the Elementary and Secondary Education Act. Pub. L. No. 107-110 (H.R. 1).

The achievement levels in this report were adopted by the Board based on a standard-setting process designed and conducted under a contract with ACT, Inc. To develop these levels, ACT convened a cross section of educators and interested citizens from across the nation and asked them to judge what students should know and be able to do relative to a body of content reflected in the NAEP framework for U.S. history. This achievement-level-setting process was reviewed by a variety of individuals including policymakers, representatives of professional organizations, teachers, parents, and other members of the general public. Prior to adopting these levels of student achievement, NAGB engaged a large number of persons to comment on the recommended levels and to review the results.

The results of the achievement-level-setting process, after NAGB's approval, became a set of achievement-level descriptions and a set of achievement-level cut points on the 0-500 NAEP U.S. history scale. The cut points are the scores that define the boundaries between below *Basic*, *Basic*, *Proficient*, and *Advanced* performance at grades 4, 8, and 12. The Board established these U.S. history achievement levels based upon the U.S. history content framework.

Achievement-Level Descriptions for Each Grade

Specific definitions of the *Basic*, *Proficient*, and *Advanced* U.S. history achievement levels for grades 4, 8, and 12 are presented in figures 1.3 through 1.5. As noted previously, the achievement levels are cumulative. Therefore, students performing at the *Proficient* level also display the competencies associated with the *Basic* level, and students at the *Advanced* level also demonstrate the skills and knowledge associated with both the *Basic* and the *Proficient* levels. For each achievement level listed in figures 1.3 through 1.5, the scale score that corresponds to the beginning of that level is shown in parentheses. For example, in figure 1.3 the scale score of 243 corresponds to the beginning of the grade 4 *Proficient* level of achievement.

Figure 1.3 Achievement Levels	Descriptions of NAEP U.S. history achievement levels for grade 4
Basic (195)	Fourth-grade students performing at the *Basic* level should be able to identify and describe a few of the most familiar people, places, events, ideas, and documents in American history. They should be able to explain the reasons for celebrating most national holidays, have some familiarity with the geography of their own state and the United States, and be able to express in writing a few ideas about a familiar theme in American history.
Proficient (243)	Fourth-grade students performing at the *Proficient* level should be able to identify, describe and comment on the significance of many historical people, places, ideas, events, and documents. They should interpret information from a variety of sources, including texts, maps, pictures, and timelines. They should be able to construct a simple timeline from data. These students should recognize the role of invention and technological change in history. They should also recognize the ways in which geographic and environmental factors have influenced life and work.
Advanced (276)	Fourth-grade students performing at the *Advanced* level should have a beginning understanding of the relationship between people, places, ideas, events, and documents. They should know where to look for information, including reference books, maps, local museums, interviews with family and neighbors, and other sources. They should be able to use historical themes to organize and interpret historical topics, and to incorporate insights from beyond the classroom into their understanding of history. These students should understand and explain the role of invention and technological change in history. They should also understand and explain the ways in which geographic and environmental factors have influenced life and work.

SOURCE: National Assessment Governing Board. U.S. History Framework for the 1994 and 2001 National Assessment of Educational Progress.

BEST COPY AVAILABLE

Figure 1.4

Descriptions of NAEP U.S. history achievement levels for grade 8

Achievement Levels	
Basic (252)	Eighth-grade students performing at the *Basic* level should be able to identify and place in context a range of historical people, places, events, ideas, and documents. They should be able to distinguish between primary and secondary sources. They should have a beginning understanding of the diversity of the American people and the ways in which people from a wide variety of national and cultural heritages have become part of a single nation. Eighth-grade students at the *Basic* level should also have a beginning understanding of the fundamental political ideas and institutions of American life and their historical origins. They should be able to explain the significance of some major historical events.
Proficient (294)	Eighth-grade students performing at the *Proficient* level should be able to explain the significance of people, places, events, ideas, and documents, and to recognize the connection between people and events within historical contexts. They should understand and be able to explain the opportunities, perspectives and challenges associated with a diverse cultural population. They should incorporate geographic, technological, and other considerations in their understanding of events and should have knowledge of significant political ideas and institutions. They should be able to communicate ideas about historical themes while citing evidence from primary and secondary sources to support their conclusions.
Advanced (327)	Eighth-grade students performing at the *Advanced* level should recognize significant themes and movements in history and begin to understand particular events in light of these themes and movements. They should have an awareness of continuity and change over time and be able to draw relevant analogies between past events and present-day situations. They should be able to frame questions about historical topics and use multiple sources to develop historical generalizations and interpretations. They should be able to explain the importance of historical themes, including some awareness of their political, social, and economic dimensions.

SOURCE: National Assessment Governing Board. U.S. History Framework for the 1994 and 2001 National Assessment of Educational Progress.

BEST COPY AVAILABLE

Figure 1.5 Descriptions of NAEP U.S. history achievement levels for grade 12

Achievement Levels	
Basic (294)	Twelfth-grade students performing at the *Basic* level should be able to identify the significance of many people, places, events, dates, ideas, and documents in U.S. history. They should also recognize the importance of unity and diversity in the social and cultural history of the United States, and an awareness of American's changing relationships with the rest of the world. They should have a sense of continuity and change in history and be able to relate relevant experience from the past to their understanding of contemporary issues. They should recognize that history is subject to interpretation and should understand the role of evidence in making an historical argument.
Proficient (325)	Twelfth-grade students performing at the *Proficient* level should understand particular people, places, events, ideas, and documents in historical context, with some awareness of the political, economic, geographic, social, religious, technological, and ideological factors that shape historical settings. They should be able to communicate reasoned interpretations of past events, using historical evidence effectively to support their positions. Their written arguments should reflect some in-depth grasp of issues and refer to both primary and secondary sources.
Advanced (355)	Twelfth-grade students achieving at the *Advanced* level should demonstrate a comprehensive understanding of events and sources of U.S. history. Recognizing that history is subject to interpretation, they should be able to evaluate historical claims critically in light of the evidence. They should understand that important issues and themes have been addressed differently at different times and that America's political, social, and cultural traditions have changed over time. They should be able to write well-reasoned arguments on complex historical topics and draw upon a wide range of sources to inform their conclusions.

SOURCE: National Assessment Governing Board. U.S. History Framework for the 1994 and 2001 National Assessment of Educational Progress.

The Trial Status of Achievement Levels

The 2001 NAEP reauthorization law requires that the achievement levels be used on a trial basis until the Commissioner of Education Statistics determines that the achievement levels are "reasonable, valid, and informative to the public."[6] Until that determination is made, the law requires the Commissioner and the Board to state clearly the trial status of the achievement levels in all NAEP reports.

In 1993, the first of several congressionally mandated evaluations of the achievement level setting process concluded that the procedures used to set the achievement levels were flawed and that the percentage of students at or above any particular achievement level cutpoint may be underestimated.[7] Others have critiqued these evaluations, asserting that the weight of the empirical evidence does not support such conclusions.[8]

In response to the evaluations and critiques, NAGB conducted an additional study of the 1992 reading achievement levels before deciding to use those reading achievement levels for reporting 1994 NAEP results.[9] When reviewing the findings of this study, the National Academy of Education (NAE) Panel expressed concern about what it saw as a "confirmatory bias" in the study and about the inability of this study to "address the panel's perception that the levels had been set too high."[10] In 1997, the NAE Panel summarized its concerns with interpreting NAEP results based on the achievement levels as follows:

First, the potential instability of the levels may interfere with the accurate portrayal of trends. Second, the perception that few American students are attaining the higher standards we have set for them may deflect attention to the wrong aspects of education reform. The public has indicated its interest in benchmarking against international standards, yet it is noteworthy that when American students performed very well on a 1991 international reading assessment, these results were discounted because they were contradicted by poor performance against the possibly flawed NAEP reading achievement levels in the following year.[11]

---

6 No Child Left Behind Act of 2001: Reauthorization of the Elementary and Secondary Education Act. Pub. L. No. 107-110 (H.R. 1).

7 United States General Accounting Office. (1993). *Education achievement standards: NAGB's approach yields misleading interpretations.* U.S. General Accounting Office Report to Congressional Requestors. Washington, DC: Author.

National Academy of Education. (1993). *Setting performance standards for achievement: A report of the National Academy of Education Panel on the evaluations of the NAEP Trial State Assessment: An evaluation of the 1992 achievement levels.* Stanford, CA: Author.

8 Cizek, G. (1993). *Reactions to National Academy of Education report.* Washington, DC: National Assessment Governing Board.

Kane, M. (1993). *Comments on the NAE evaluation of the NAGB achievement levels.* Washington, DC: National Assessment Governing Board.

9 American College Testing. (1995). *NAEP reading revisited: An evaluation of the 1992 achievement level descriptions.* Washington, DC: National Assessment Governing Board.

10 National Academy of Education. (1996). Reading achievement levels. In *Quality and utility: The 1994 Trial State Assessment in reading. The fourth report of the National Academy of Education Panel on the evaluation of the NAEP Trial State Assessment.* Stanford, CA: Author.

11 National Academy of Education. (1997). *Assessment in transition: Monitoring the nation's educational progress* (p. 99). Mountain View, CA: Author.

BEST COPY AVAILABLE

The National Center for Education Statistics and the National Assessment Governing Board have sought and continue to seek new and better ways to set performance standards on NAEP.[12] For example, NCES and NAGB jointly sponsored a national conference on standard setting in large-scale assessments, which explored many issues related to standard setting.[13] Although new directions were presented and discussed, a proven alternative to the current process has not yet been identified. The Deputy Commissioner of Education Statistics and the Board continue to call on the research community to assist in finding ways to improve standard setting for reporting NAEP results.

The most recent congressionally mandated evaluation conducted by the National Academy of Sciences (NAS) relied on prior studies of achievement levels, rather than carrying out new evaluations, on the grounds that the process has not changed substantially since the initial problems were identified. Instead, the NAS Panel studied the development of the 1996 science achievement levels. The NAS Panel basically concurred with earlier congressionally mandated studies. The Panel concluded that "NAEP's current achievement level setting procedures remain fundamentally flawed. The judgment tasks are difficult and confusing; raters' judgments of different item types are internally inconsistent; appropriate validity evidence for the cut scores is lacking; and the process has produced unreasonable results."[14]

The NAS Panel accepted the continuing use of achievement levels in reporting NAEP results on a developmental basis, until such time as better procedures can be developed. Specifically, the NAS Panel concluded that "....tracking changes in the percentages of students performing at or above those cut scores (or, in fact, any selected cut scores) can be of use in describing changes in student performance over time."[15]

The National Assessment Governing Board urges all who are concerned about student performance levels to recognize that the use of these achievement levels is a developing process and is subject to various interpretations. The Board and the Deputy Commissioner believe that the achievement levels are useful for reporting trends in the educational achievement of students in the United States.[16] In fact, achievement level results have been used in reports by the President of the United States, the

---

12 Reckase, Mark, D. (2000). *The evolution of the NAEP achievement levels setting process: A summary of the research and development efforts conducted by ACT.* Iowa City, IA: ACT, Inc.

13 National Assessment Governing Board and National Center for Education Statistics. (1995). *Proceedings of the joint conference on standard setting for large-scale assessments of the National Assessment Governing Board (NAGB) and the National Center for Education Statistics* (NCES). Washington, DC: Government Printing Office.

14 Pellegrino, J.W., Jones, L.R., & Mitchell, K.J. (Eds.). (1998). *Grading the nation's report card: evaluating NAEP and transforming the assessment of educational progress.* Committee on the Evaluation of National Assessments of Educational Progress, National Research Council. (p.182). Washington, DC: National Academy Press.

15 Ibid., page 176.

16 Forsyth, Robert A. (2000). A description of the standard-setting procedures used by three standardized test publishers. In *Student performance standards on the National Assessment of Educational Progress: Affirmations and improvements.* Washington, DC: National Assessment Governing Board.

Nellhaus, Jeffrey M. (2000). States with NAEP-like performance standards. In *Student performance standards on the National Assessment of Educational Progress: Affirmations and improvements.* Washington, DC: National Assessment Governing Board.

BEST COPY AVAILABLE

Secretary of Education, state governors, legislators, and members of Congress. Government leaders in the nation and in more than 40 states use these results in their annual reports.

However, based on the congressionally mandated evaluations so far, the Deputy Commissioner agrees with the National Academy's recommendation that caution needs to be exercised in the use of the current achievement levels. Therefore, the Deputy Commissioner concludes that these achievement levels should continue to be used on a trial basis and should continue to be interpreted with caution.

Interpreting NAEP Results

The average scores and percentages presented in this report are estimates because they are based on samples of students rather than on entire populations. Moreover, the collection of questions used at each grade level is but a sample of the many questions that could have been asked to assess student knowledge of the framework content. As such, the results are subject to a measure of uncertainty, reflected in the standard error of the estimates. The standard errors for the estimated scale scores and percentages in this report are provided in appendix B.

The differences between scale scores and between percentages discussed in the following chapters take into account the standard errors associated with the estimates. Comparisons are based on statistical tests that consider both the magnitude of the difference between the group average scores or percentages and the standard errors of those statistics. Throughout this report, differences between scores or between percentages are pointed out only when they are significant from a statistical perspective. All differences reported are significant at the 0.05 level with appropriate adjustments for multiple comparisons. The term significant is not intended to imply a judgment about the absolute magnitude or the educational relevance of the differences. It is intended to identify statistically dependable population differences to help inform dialogue among policymakers, educators, and the public.

Readers are cautioned against interpreting NAEP results in a causal sense. Inferences related to student subgroup performance or to the effectiveness of public and nonpublic schools, for example, should take into consideration the many socioeconomic and educational factors that may also impact on performance in U.S. history.

Overview of the Remaining Report

The results in chapters 2 and 3 of this report are based on the set of data with no accommodations offered to students. Findings are presented for the nation and for all the major reporting subgroups included in all NAEP report cards. Comparisons with results from the 1994 assessment are noted where the data permit.

NAEP has sought to assess samples that are as inclusive as possible. Nevertheless, there has always been some exclusion of students with disabilities (SD) and limited English proficient (LEP) students who could not be assessed meaningfully without accommodations. Local school officials have made decisions about exclusion in accordance with explicit criteria provided by the NAEP program. In order to expand the proportion of students who can be assessed meaningfully, the NAEP program began in recent assessments to explore the use of accommodations with special-needs students. Chapter 4 presents an overview of a second set of results—those that include students who were provided accommodations during the test administration. By including these results in the nation's U.S. history report card, the NAEP program continues a phased transition toward a more inclusive reporting sample. Future assessment results will be based solely on a student and school sample in which accommodations are permitted.

Chapter 5 provides sample assessment questions and student responses from the 2001 assessment. Also presented in chapter 5 are item maps that position selected question descriptions along the NAEP U.S. history scale where they are likely to be answered successfully by students. The descriptions used on these item maps focus on the U.S. history skill or knowledge needed to answer the question. Chapter 6 examines contexts for learning U.S. history in terms of classroom practices and student variables. The data presented in both chapters 5 and 6 are based on the set of results that did not include accommodated special-needs students.

This report also contains appendices that support or augment the results presented. Appendix A contains an overview of the NAEP U.S. history framework and specifications, information on the national sample, and a more detailed description of the major reporting subgroups featured in chapters 2 and 3. Appendix B contains the full data with standard errors for all tables and figures in this report. Appendix C contains a list of the NAEP U.S. history committee members.

2 Average Scale Scores and Achievement-Level Results

Overview

This chapter presents results for the nation from the NAEP 2001 U.S. history assessment at grades 4, 8, and 12. Student performance is described in two ways: one, by average scores on the NAEP U.S. history scale, which ranges from 0 to 500; and two, in terms of the percentages of students who attained each of the three U.S. history achievement levels: *Basic*, *Proficient*, and *Advanced*.

Chapter Focus

Are the nation's fourth-, eighth-, and twelfth-graders making progress in U.S. history?

Results of the NAEP 2001 U.S. history assessment are compared with results from the previous assessment, which took place in 1994. This comparison is possible because the assessments share a common set of tasks based on the current U.S. history framework and because the population of students in both years was sampled and assessed using comparable procedures. The results presented in this chapter are based on a representative sample of students assessed under conditions that did not permit accommodations for special-needs students. These were the same conditions under which the 1994 history assessment was administered, thus making it possible to report trends in student performance across the assessment years. A second set of results, reflecting part of a phased transition toward a more inclusive reporting sample in which accommodations were permitted for special-needs students, is presented in chapter 4.

Chapter Contents

Overview

Average Scale Score Results

Scale Scores by Percentile

Achievement-Level Results

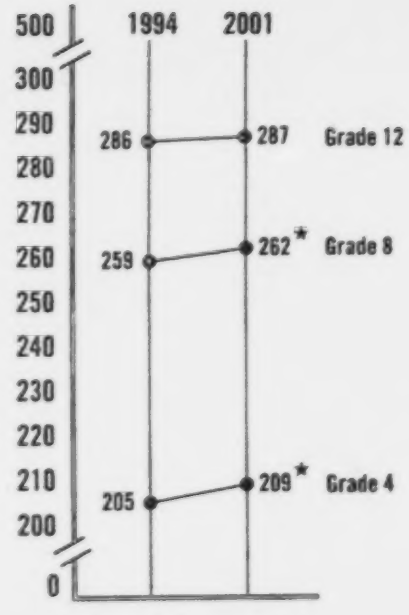
BEST COPY AVAILABLE

Average Scale Score Results

The results of the NAEP 2001 U.S. history assessment show improvement in student performance over the 1994 results at grades 4 and 8, but do not show a statistically significant change at grade 12. Figure 2.1 presents the average U.S. history scale scores for the nation for fourth-, eighth-, and twelfth-graders attending both public and nonpublic schools in 1994 and 2001.

Figure 2.1 Average U.S. history scale scores, grades 4, 8, and 12: 1994 and 2001

National Scale Score Results



★ Significantly different from 1994.

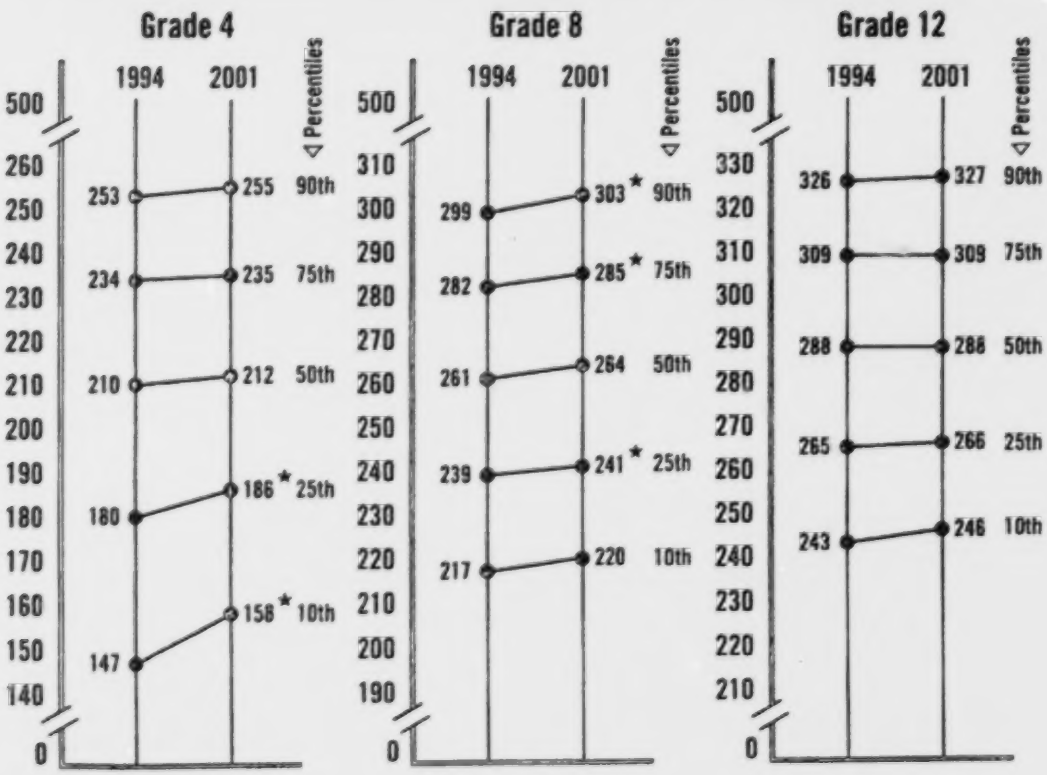
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

Scale Scores by Percentile

One way to view students' scale score performance on the U.S. history assessment is by looking at the scale scores attained by students across the performance distribution. The percentile indicates the percentage of students whose scores fell below a particular average score. The advantage of looking at the data this way is that it reveals changes in performance for both lower- and higher-performing students. Figure 2.2 presents the U.S. history scale scores for grades 4, 8, and 12 at the 10th, 25th, 50th, 75th, and 90th percentiles for both the 1994 and the 2001 assessments.

Figure 2.2 U. S. history scale score percentiles, grades 4, 8, and 12: 1994 and 2001

National Performance Distribution



* Significantly different from 1994.

SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

At grade 4, the scale scores at the lower percentiles (10th and 25th) were higher in 2001 than in 1994. At the other percentiles for grade 4, apparent changes since 1994 were not statistically significant. Increases in average scores at grade 8 were evident among students at both the lower percentile (25th) and the upper percentiles (75th and 90th). At the 10th and 50th percentiles, however, apparent changes since 1994 were not statistically significant. There were no statistically significant changes in average scores when viewed across the score distribution at grade 12.

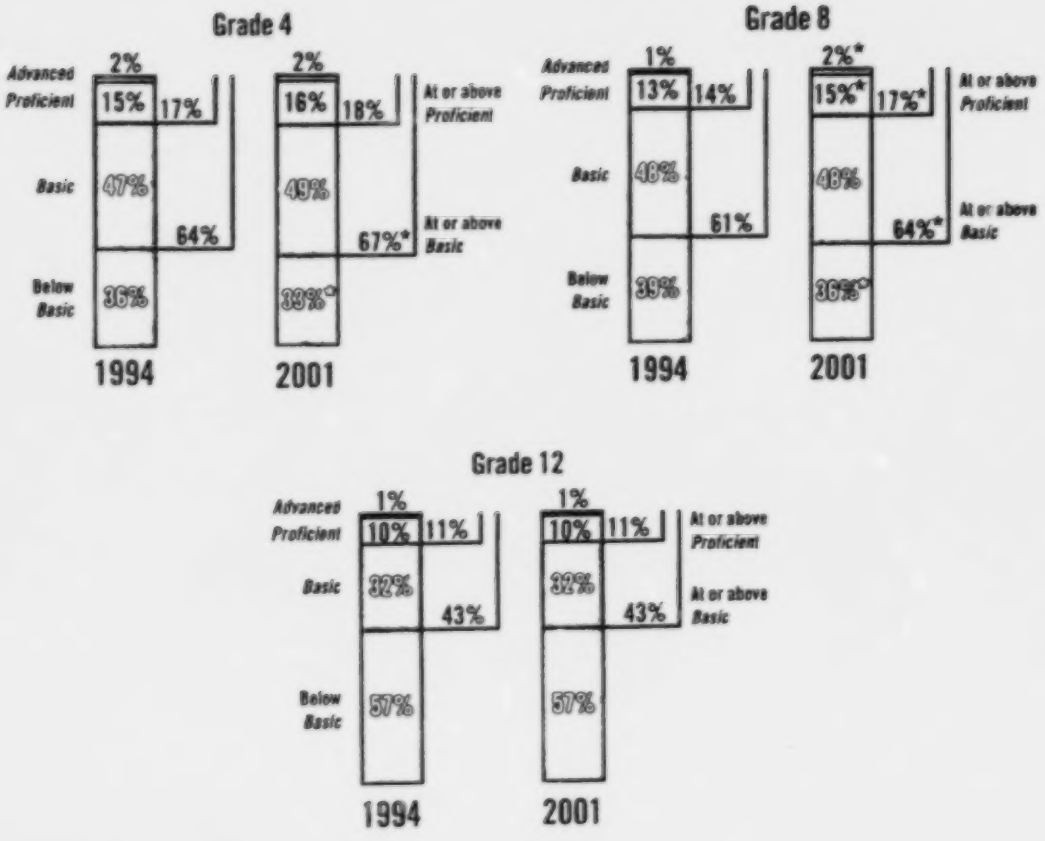
Achievement-Level Results

The achievement levels are performance standards adopted by the National Assessment Governing Board, based on the collective judgements of experts about what students should be expected to know and be able to do in terms of the NAEP U.S. history framework. A discussion of the trial status of achievement levels is in chapter 1.

Achievement-level results for each grade are presented in figure 2.3. Results are presented in two ways: 1) the percentage of students within each achievement-level range, and 2) the percentage of students at or above the *Basic* and at or above the *Proficient* levels. In reading figure 2.3, it is necessary to keep in mind that the percentages at or above specific achievement levels are cumulative. For example, included among the percentage of students at or above the *Basic* level are also those who have achieved the *Proficient* and *Advanced* levels of performance.

Figure 2.3
National Achievement Level Results

Percentage of students within and at or above U.S. history achievement levels, grades 4, 8, and 12: 1994 and 2001



★ Significantly different from 1994.
NOTE: Percentages within each U.S. history achievement level may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

BEST COPY AVAILABLE

In the 2001 U.S. history assessment, 18 percent of fourth-graders, 17 percent of eighth-graders, and 11 percent of twelfth-graders performed at or above the *Proficient* level—identified by NAGB as the level at which all students should perform. Students' attainment of the achievement levels across years generally reflects the changes in scale score results described in the previous section: improvement at some levels at grade 4, general improvement at grade 8, and no statistically significant change at grade 12.

At grade 4, the percentage of students performing at or above *Basic* in 2001 was higher than in 1994. At grade 8, there was improvement across the levels compared to the previous assessment, with the percentages of students performing at or above *Basic*, at or above *Proficient*, and at *Advanced* increasing from 1994 to 2001. At grade 12, however, there was no statistically significant change in the percentages of students performing at each level. As in 1994, only a small percentage of students at each grade performed at the *Advanced* level, with 2 percent at grades 4 and 8, and 1 percent at grade 12 attaining that level. Fifty-seven percent of twelfth-graders were below the *Basic* achievement level in 2001.

3 Average Scale Scores and Achievement-Level Results for Selected Subgroups

In addition to reporting on the performance of all students, NAEP also provides results for various subgroups of students at each grade. Examining subgroup results provides insight, not only into how these groups of students performed in comparison to one another, but also into how each group has progressed over time. The information presented in this chapter serves as a valuable indicator of the progress of subgroups of the students across the nation.

Chapter Focus

Are selected subgroups of students making progress in U.S. history?

Results for the NAEP 2001 U.S. history assessment are presented by gender, race/ethnicity, region of the country, parents' highest level of education, type of school, type of location, and eligibility for the Free/Reduced-Price School Lunch program. For all subgroups except two—type of location and free/reduced-price school lunch eligibility—results are available from 1994, and are presented here for comparison with results from 2001.

All differences reported in this chapter between demographic subgroups for the 2001 assessment and between the 2001 and the 1994 results are based on statistical tests that consider both the magnitude of the difference between the group average scores or percentages and the standard errors of those statistics. Differences between groups and between assessment years are discussed only if they have been determined to be statistically significant. Furthermore, the reader should bear in mind that

Chapter Contents

Gender

Race/Ethnicity

Region of the Country

Parents' Education

Type of School

Type of Location

Eligibility for Free/Reduced-Price School Lunch Program

BEST COPY AVAILABLE

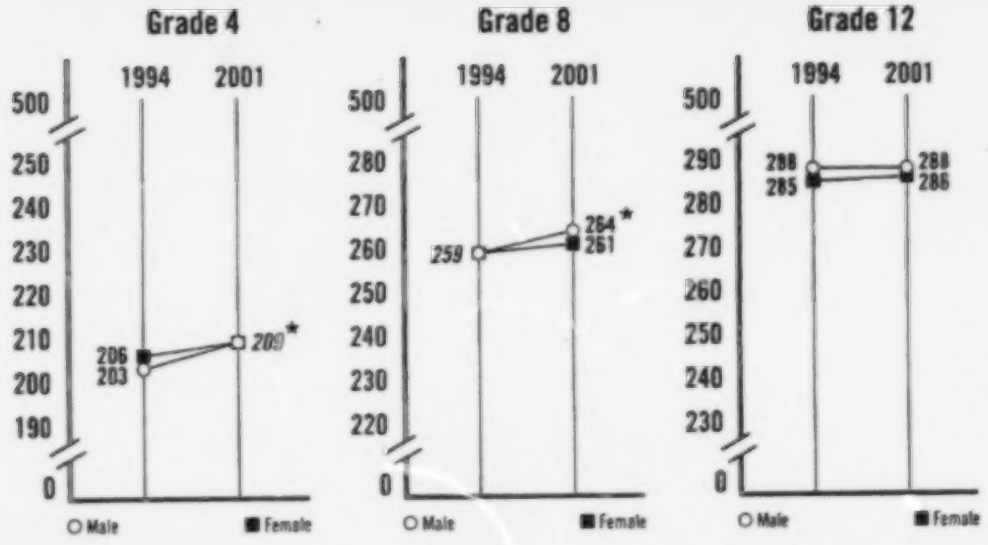
differences in performance among subgroups of students most likely reflect a range of socioeconomic and educational factors not addressed in this report or by NAEP.

Gender

Figure 3.1 presents the 1994 and 2001 average U.S. history scale scores by gender. Patterns in performance among male and female students varied somewhat by grade. At grade 4 both male and female students had higher average scores in 2001 than in 1994, while at grade 8 only males showed a gain in 2001 over 1994. The apparent gain for female eighth-graders was not statistically significant. At grade 12, there was no statistically significant change in the performance of male and female students from one assessment to the next. At all three grades in 2001, there was no statistically significant difference between the performance of males and females.

Figure 3.1 Average U.S. history scale scores by gender, grades 4, 8, and 12: 1994 and 2001

National Scale Score Results by Gender



* Significantly different from 1994.

NOTE: Italicized scale score values indicate that two or more groups had the same rounded average score. The average scale scores, when rounded, were the same for male and female students at grade 4 in 2001 and grade 8 in 1994.

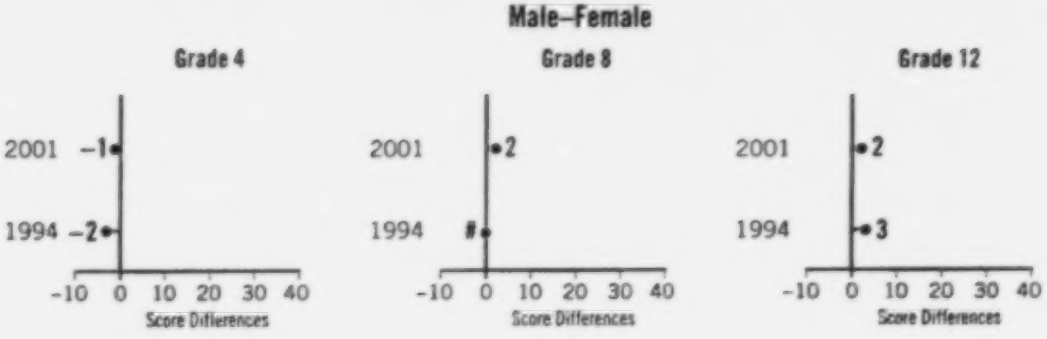
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

BEST COPY AVAILABLE

The performance of subgroups on the U.S. history assessment can also be compared by determining whether a difference or "gap" exists between groups' average scores and, if it does, whether that gap increases or decreases between assessment years. As stated previously, no significant difference between male or female average scores was observed in 2001. Figure 3.2 also shows that there was no statistically significant change between 1994 and 2001 in these small and nonsignificant gender gaps.

Figure 3.2
National Scale Score Differences by Gender

Differences in average U.S. history scale scores by gender, grades 4, 8, and 12: 1994 and 2001



Difference is between -0.5 and 0.5.
NOTE: Score differences are calculated based on differences between unrounded average scale scores.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

The percentages of male and female students at or above the U.S. history achievement levels and within each achievement level range are presented in figure 3.3. Achievement-level results for males and females at both grades 4 and 12 showed no statistically significant increases or decreases since 1994. At grade 8, however, the percentages of male students at or above *Basic*, at or above *Proficient*, and at *Advanced* were all higher in 2001 than in 1994, while any apparent changes in the percentages of females at or above any of the achievement levels were not statistically significant. A comparison of the percentages of male and female students at or above the *Basic* and *Proficient* levels in 2001 shows no difference at grade 4, but does show a higher percentage of males than females at or above *Proficient* at grade 8, and a higher percentage of males than females at or above *Basic* at grade 12.

Figure 3.3
National Achievement-Level Results by Gender

Percentage of students within and at or above U.S. history achievement levels by gender, grades 4, 8, and 12: 1994 and 2001

* Significantly different from 1994.
NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

BEST COPY AVAILABLE

Race/Ethnicity

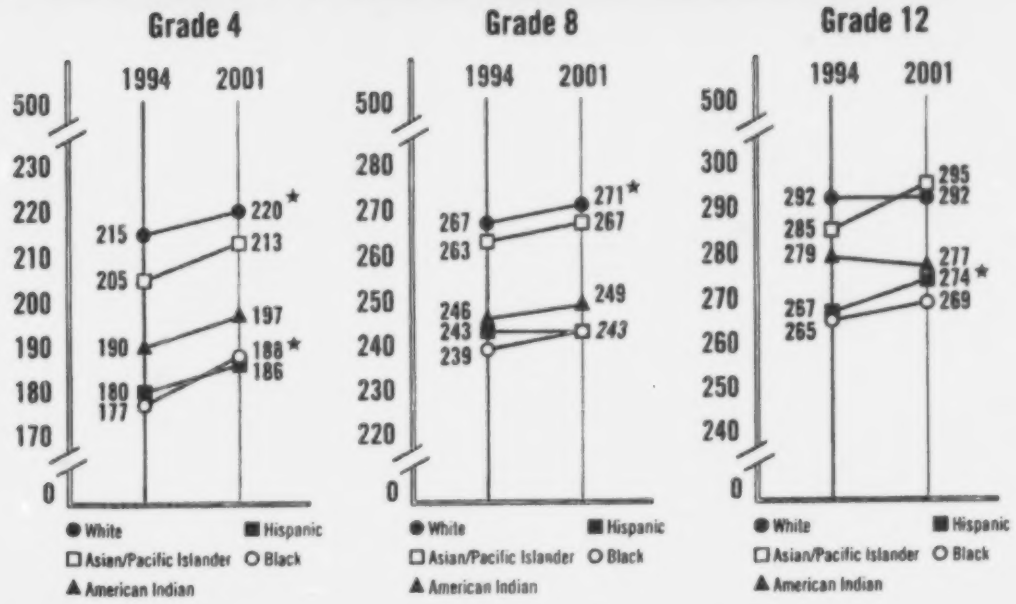
Students participating in the U.S. history assessment were asked to indicate which of the following racial/ethnic subgroups best described them—White, Black, Hispanic, Asian/Pacific Islander, or American Indian (including Alaska Native). Figure 3.4 presents average scale scores for students by these subgroups at grades 4, 8, and 12.

At grade 4, both White students and Black students had higher average scores in 2001 than in 1994. At grade 8, White students showed a gain since 1994, and at grade 12 Hispanic students had higher average scores in 2001, compared to 1994. No other changes were statistically significant.

In 2001, differences in performance by racial/ethnic subgroup continue to be evident at all three grades. On average, White students had higher scores than Black, Hispanic, and American Indian students at all three grades. Asian/Pacific Islander students had higher average scores than Black and Hispanic students at all three grades. White fourth-grade students had higher average scores than Asian/Pacific Islander fourth-graders. These differences should be interpreted with caution. The average score of a selected subgroup does not represent the entire range of performance within that group. Furthermore, differences between groups of students cannot be attributed solely to group identification. A complex array of educational and social factors interacts to affect average student performance.

Figure 3.4
National Scale Score Results by Race/Ethnicity

Average U.S. history scale scores by race/ethnicity, grades 4, 8, and 12: 1994 and 2001



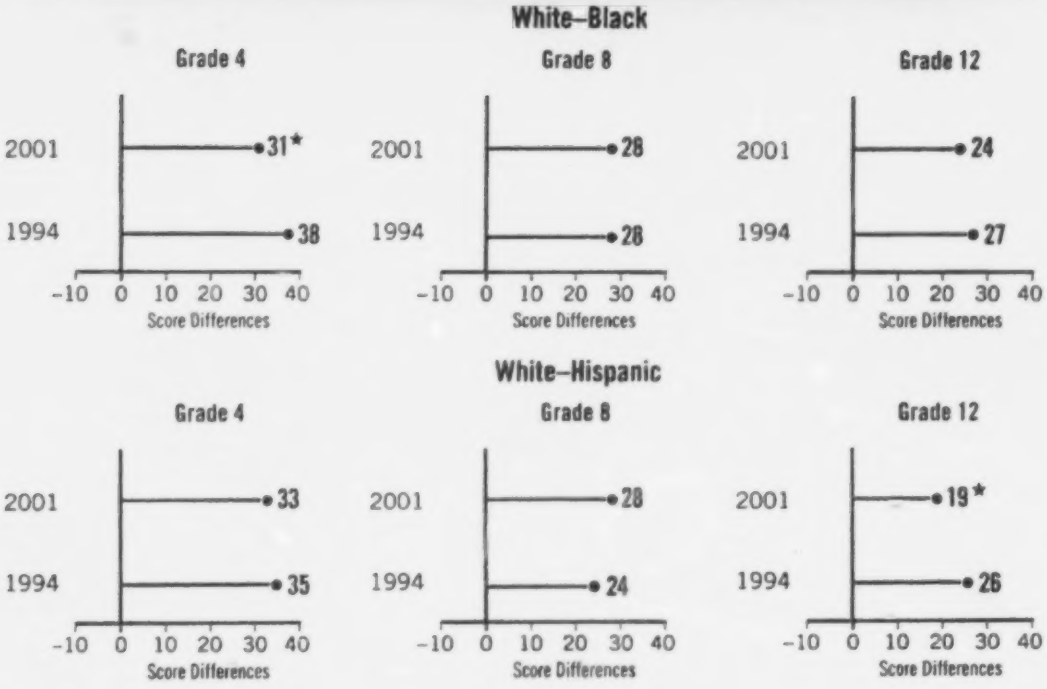
★ Significantly different from 1994.
NOTE: Italicized scale score values indicate that two or more groups had the same rounded average score. The average scale scores, when rounded, were the same for Hispanic and Black students at grade 8 in 2001.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

BEST COPY AVAILABLE

Scale score differences between White students and Black students and between White students and Hispanic students are presented in figure 3.5. Results from the 2001 U.S. history assessment reflect a narrowing of the score gap between White students and Black students at grade 4, and between White students and Hispanic students at grade 12.

Figure 3.5
National Scale Score Differences by Race/Ethnicity

Differences in average U.S. history scale scores by race/ethnicity, grades 4, 8, and 12: 1994 and 2001



* Significantly different from 1994.
NOTE: Score differences are calculated based on differences between unrounded average scale scores.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

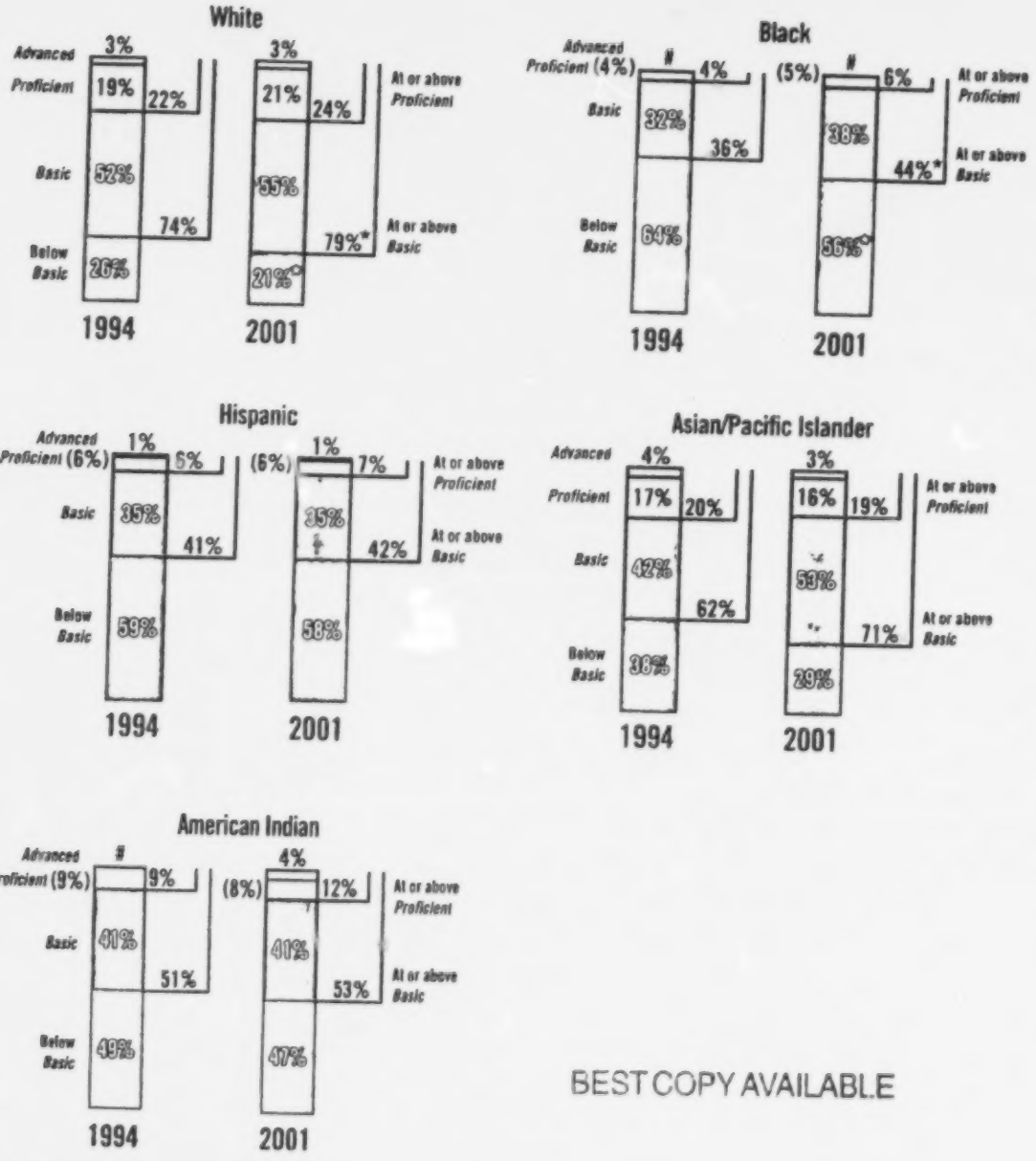
Achievement-level results for the racial/ethnic subgroups are presented in figure 3.6a, b, and c. While there have been some overall achievement gains since 1994 at grades 4 and 8, not all groups show improvement. At grade 4, both White students and Black students had higher percentages at or above *Basic* in 2001 compared to 1994. At grade 8, however, only White students showed an increase in the percentages at or above *Proficient* and at *Advanced*. At grade 12, none of the apparent changes between 1994 and 2001 in the percentages of students at or above any of the history achievement levels were statistically significant.

Comparing the subgroups' performance in 2001 shows higher percentages of White students and Asian/Pacific Islander students at or above the *Basic* and *Proficient* levels than Black and Hispanic students at all three grades.

Figure 3.6a

National Achievement-Level Results by Race/Ethnicity

Percentage of students within and at or above U.S. history achievement levels by race/ethnicity, grade 4: 1994 and 2001



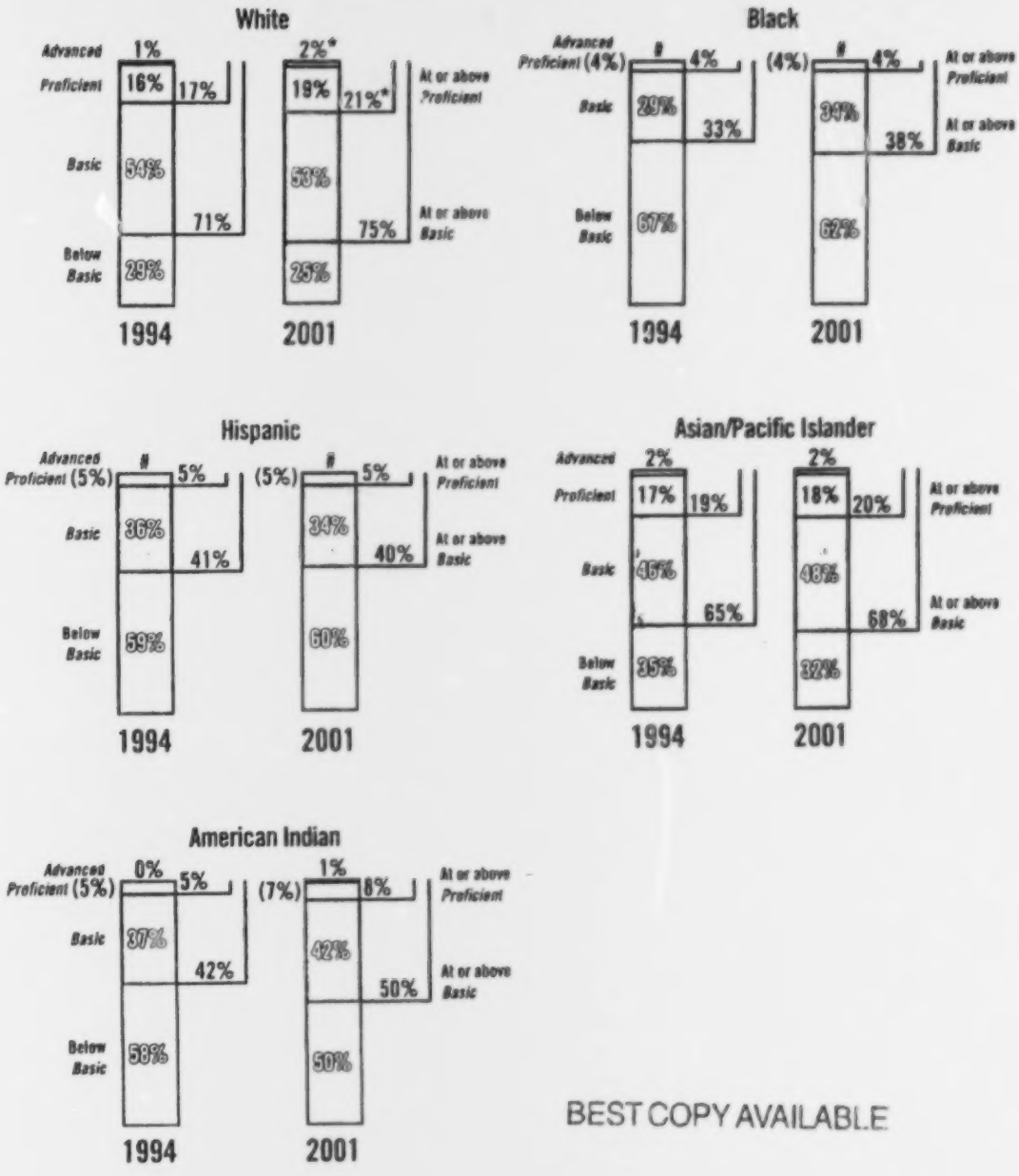
BEST COPY AVAILABLE

★ Significantly different from 1994.
Percentage is between 0.0 and 0.5.
NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

Figure 3.6b

National Achievement Level Results by Race/Ethnicity

Percentage of students within and at or above U.S. history achievement levels by race/ethnicity, grade 8: 1994 and 2001



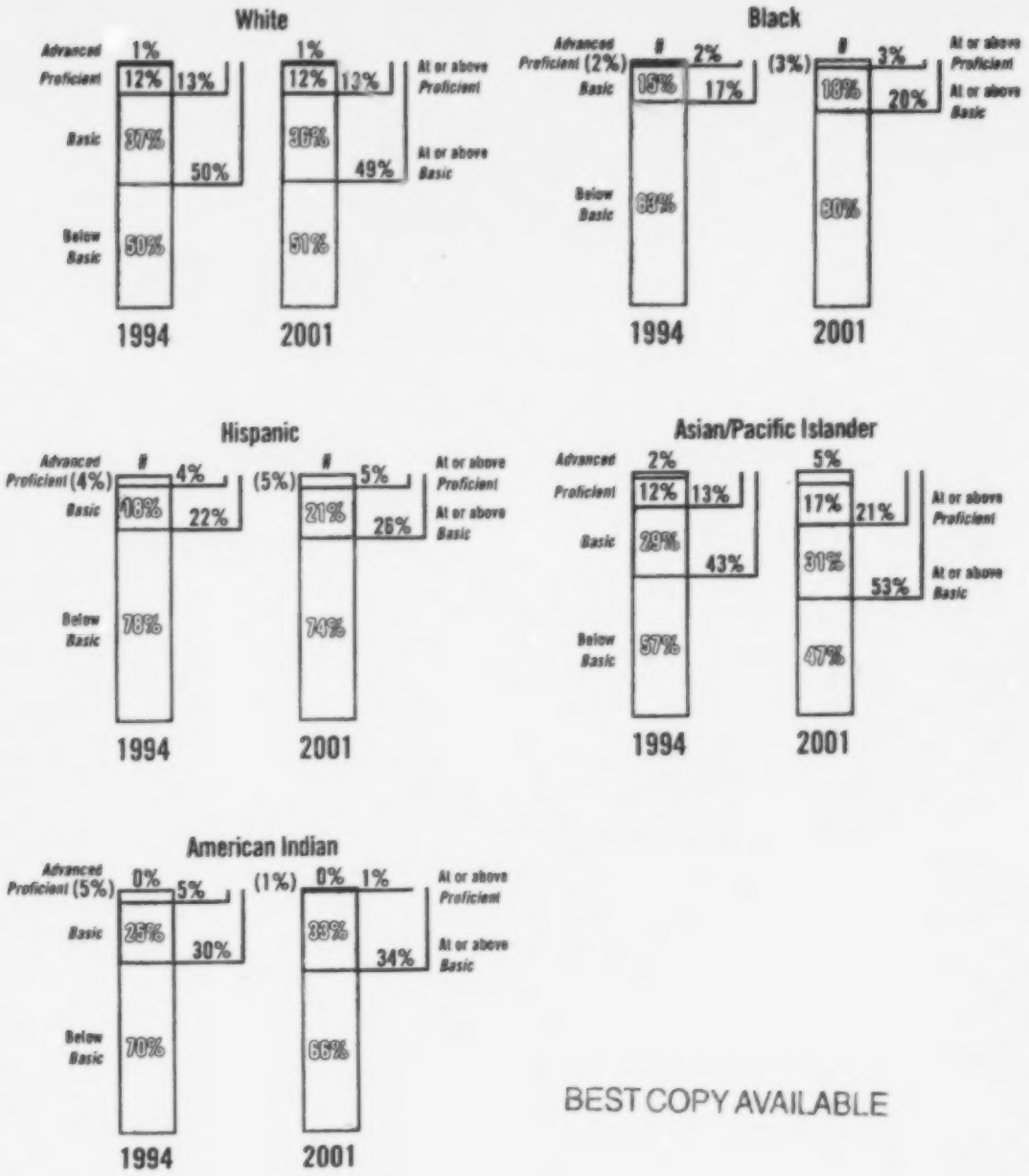
BEST COPY AVAILABLE

★ Significantly different from 1994.
Percentage is between 0.0 and 0.5.
NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

Figure 3.6c

National Achievement-Level Results by Race/Ethnicity

Percentage of students within and at or above U.S. history achievement levels by race/ethnicity, grade 12: 1994 and 2001



BEST COPY AVAILABLE

Percentage is between 0.0 and 0.5

NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.

SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

Region of the Country

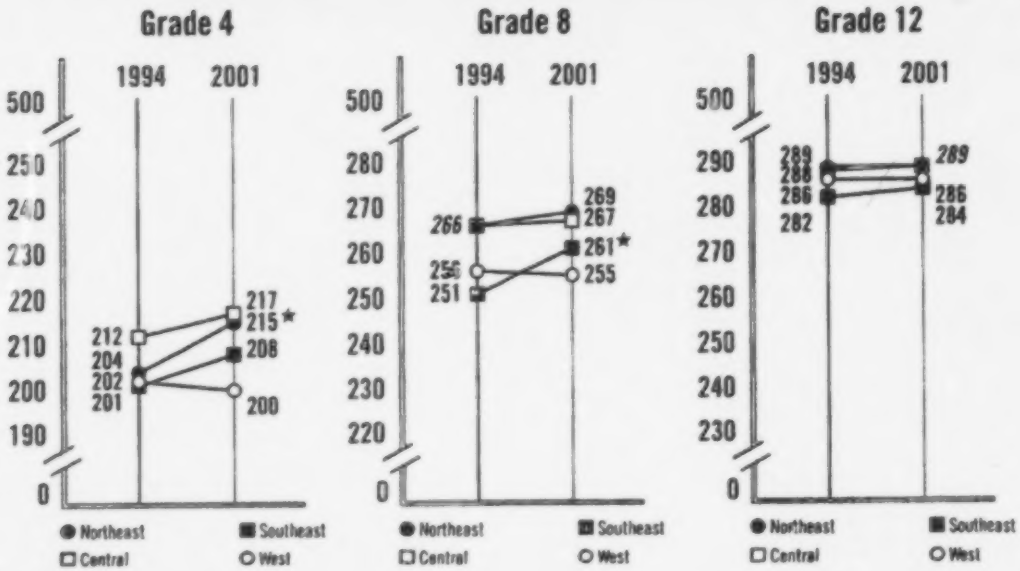
NAEP assessments traditionally provide results for four regions of the country: Northeast, Southeast, Central, and West. Appendix A (see page 129) contains a description of the states and other jurisdictions that make up each region.

Scale score results by region are presented in figure 3.7. Although overall gains in student performance were observed at grades 4 and 8, not all regions showed increases. At grade 4, only the Northeast region showed a gain in the U.S. history average score since 1994, while at grade 8, the only increase occurred in the Southeast region. None of the other apparent changes between 1994 and 2001 in regional average scores were statistically significant. Regional results at grade 12 were consistent with the overall national results that did not show a change in students' performance from 1994 to 2001.

Some differences in performance between regions of the country were evident in 2001. At both grades 4 and 8, students in the Northeast, Southeast, and Central regions all had higher average scores than students in the West, and students in the Central region outperformed their peers in the Southeast on average. At grade 12, none of the apparent differences in average scores among the four regions were statistically significant.

Figure 3.7
National Scale Score Results by Region of the Country

Average U.S. history scale scores by region of the country, grades 4, 8, and 12: 1994 and 2001



★ Significantly different from 1994.
NOTE: Italicized scale score values indicate that two or more groups had the same rounded average score. The average scale scores, when rounded, were the same for Northeast and Central regions at grade 8 in 1994 and at grade 12 in 2001.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

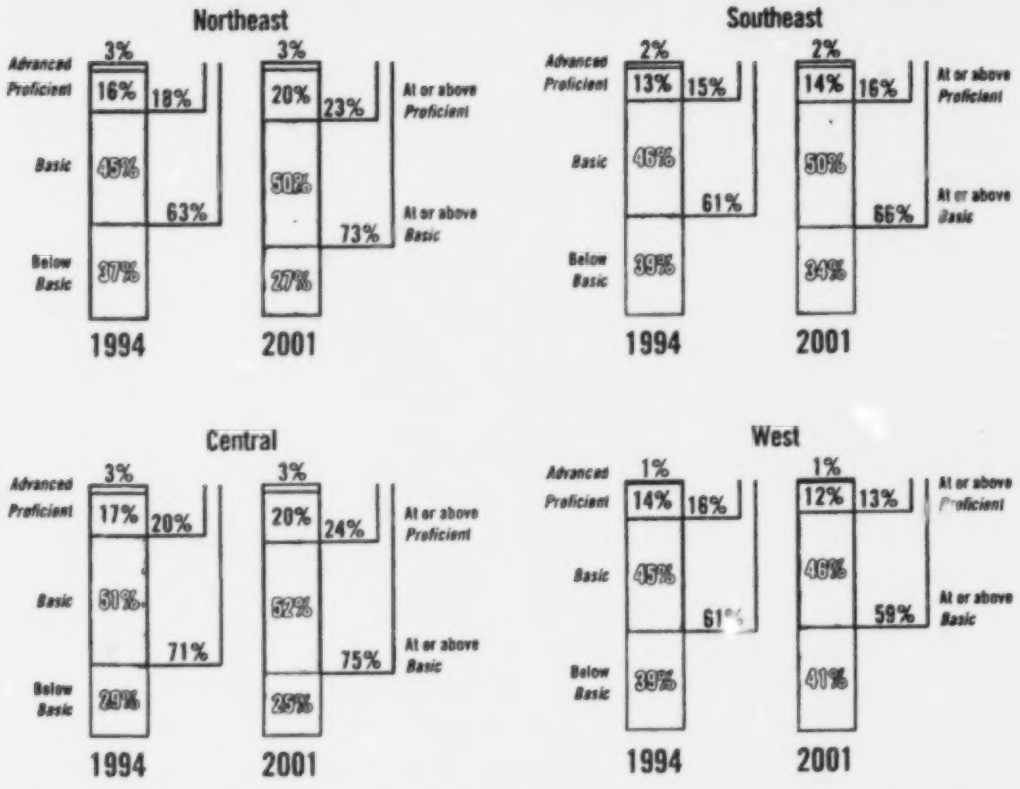
BEST COPY AVAILABLE

Achievement-level results for the four regions are displayed in figure 3.8a, b, and c by grade. The only gains occurred in the Southeast region at grade 8, where the percentage both at or above *Basic* and at or above *Proficient* increased in 2001 compared to 1994.

As with the scale score results by region, some differences between regions in the percentages of students at or above the different achievement levels were evident in 2001. A higher percentage of fourth-grade students in the Northeast and Central regions were at or above both the *Basic* and *Proficient* levels than in the West, and a higher percentage of fourth-graders in the Central region were at or above *Basic* than in the Southeast. There were also higher percentages of eighth-grade students in the Northeast and Central regions at or above *Basic* and *Proficient* than in the West. In addition, the percentage of eighth-graders at or above *Basic* was higher in the Northeast and Central region than in the Southeast, and higher in the Southeast than in the West.

Figure 3.8a
National Achievement Level Results by Region of the Country

Percentage of students within and at or above U.S. history achievement levels by region of the country, grade 4: 1994 and 2001

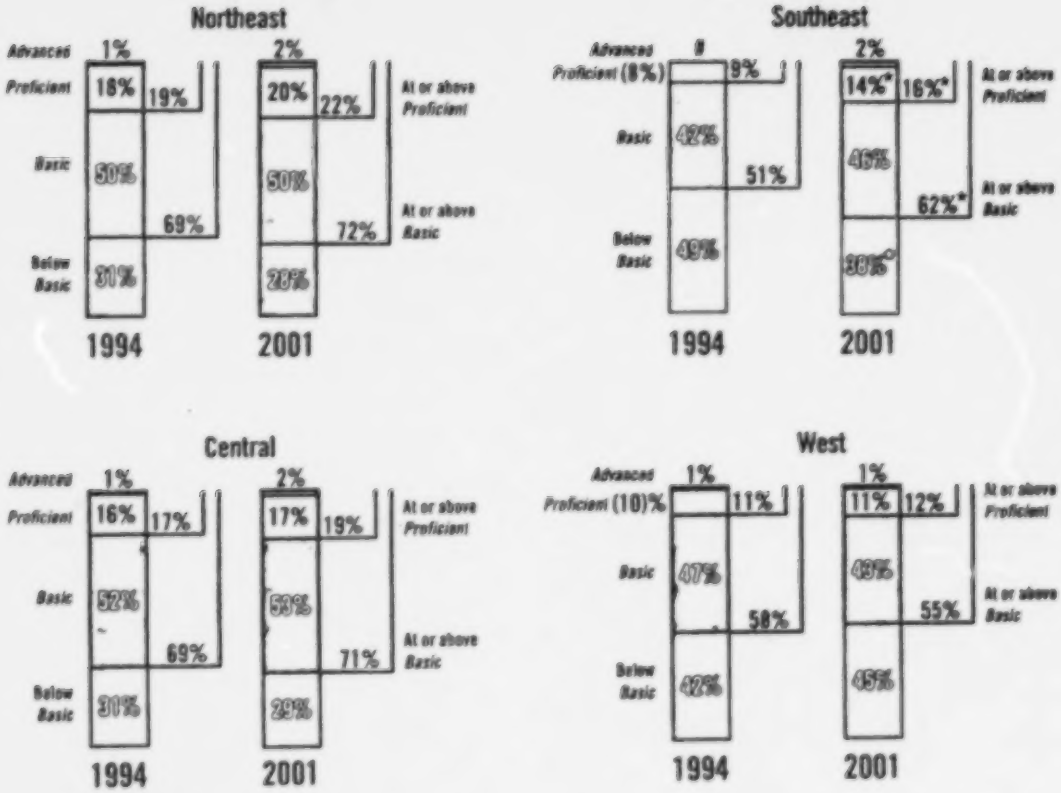


NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

Figure 3.8b

National Achievement-Level Results by Region of the Country

Percentage of students within and at or above U.S. history achievement levels by region of the country, grade 8: 1994 and 2001

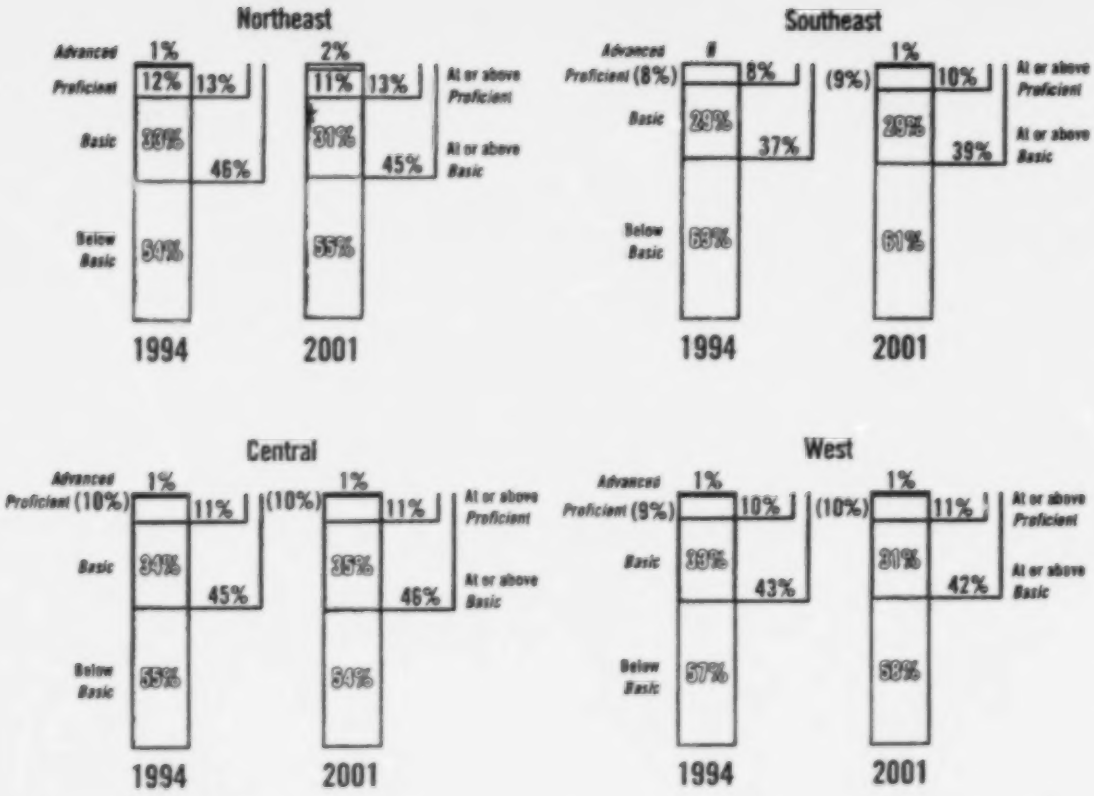


★ Significantly different from 1994.
Percentage is between 0.0 and 0.5.
NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

Figure 3.8c

National Achievement Level Results by Region of the Country

Percentage of students within and at or above U.S. history achievement levels by region of the country, grade 12 1994 and 2001



Percentage is between 0.0 and 0.5.

NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.

SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

Parents' Highest Level of Education

Eighth- and twelfth-grade students who participated in the NAEP U.S. history assessment were asked to indicate the highest level of education completed by each parent. Four levels of education were identified: did not finish high school, graduated from high school, some education after high school, and graduated from college. Students could also choose the response, "I don't know." For this analysis, the highest education level reported for either parent was used. Data were not collected at grade 4 because in previous NAEP assessments fourth-graders' responses about their parents' education were highly variable and contained a large percentage of "I don't know" responses.

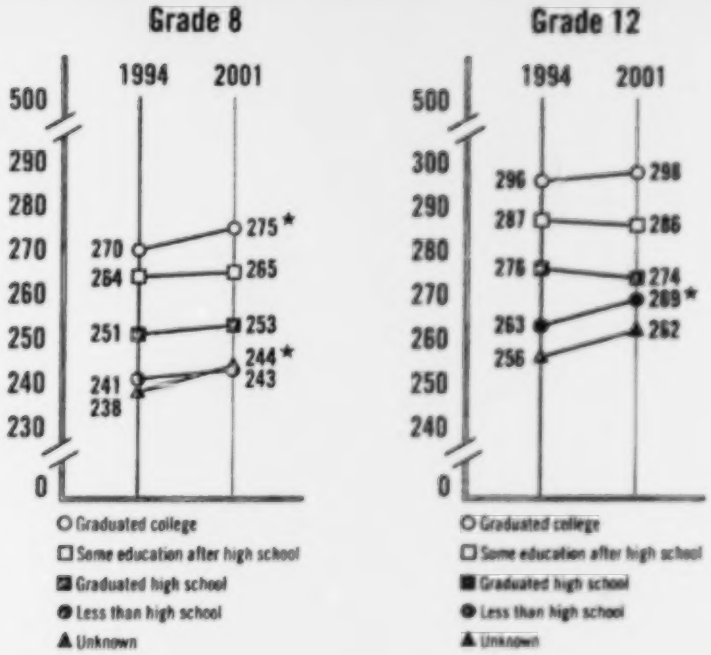
The scale score results for all levels of student-reported parent education level are presented in figure 3.9. In 2001, almost one-half of the eighth- and twelfth-grade students reported at least one parent had graduated from college (47 and 46 percent, respectively), whereas only a small percentage reported that their parents had not graduated high school (7 percent at both grades). Additional information on the percentage of students reporting parents' highest level of education is available in appendix B.

At grade 8, students who reported that at least one parent graduated from college had higher average scores in 2001 than in 1994. At grade 12, there was an increase in the average scores of students who reported that neither parent finished high school. The 2001 results indicate that, overall, there was a clear positive relationship between parent education level and the performance of both eighth- and twelfth-graders on the U.S. history assessment. At both grades, the higher the parental education level reported, the higher the average score attained.

Figure 3.9

National Scale Score Results by Parents' Education

Average U.S. history scale scores by parents' highest level of education, grades 8 and 12: 1994 and 2001



★ Significantly different from 1994.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

Achievement-level results across years by level of parental education are presented in figure 3.10a and b.

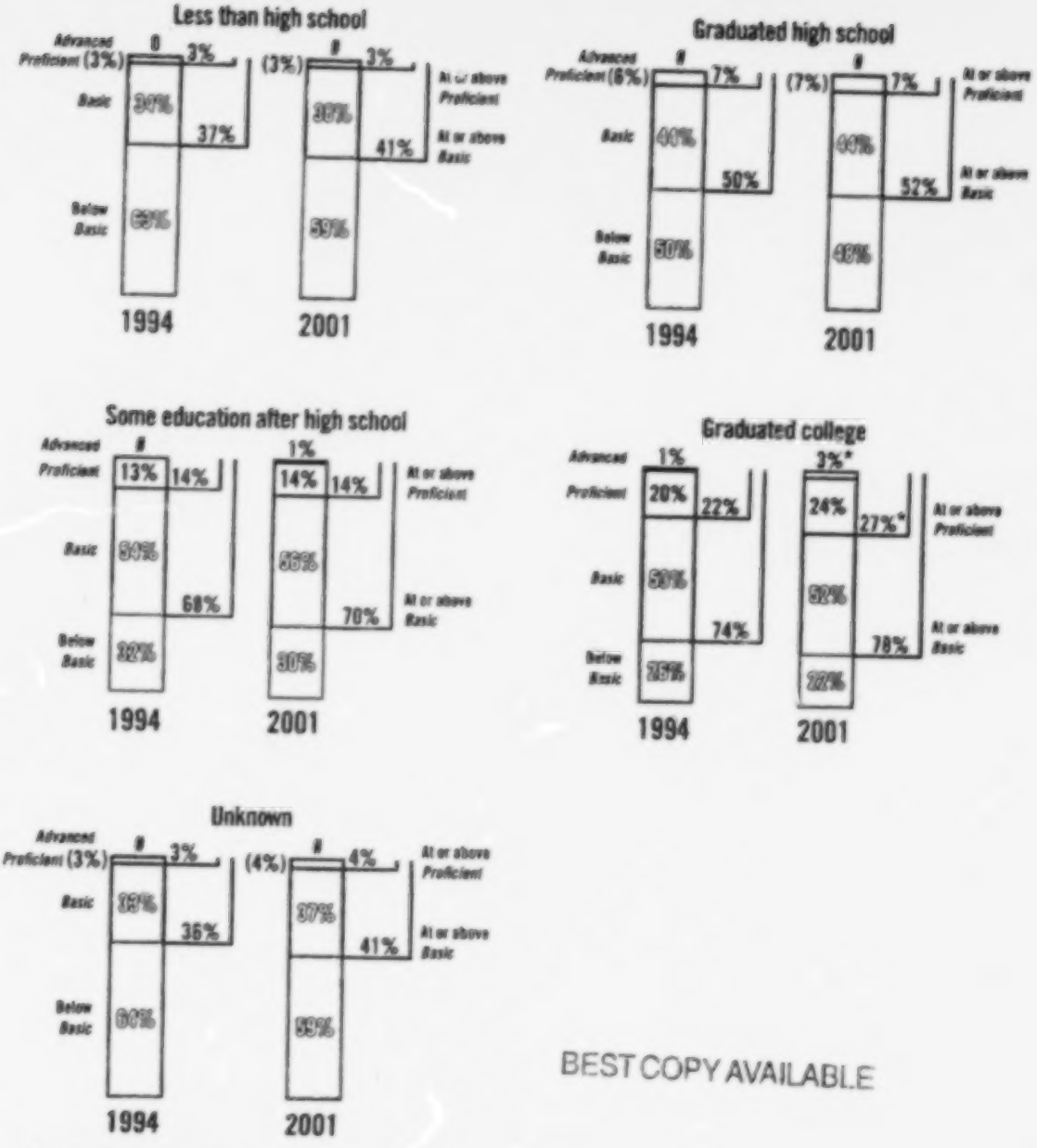
There were higher percentages of eighth-graders at or above *Proficient* and at *Advanced* in 2001 among students who reported that at least one parent graduated from college. None of the other changes in achievement level results by level of parental education were statistically significant.

As with the average scale score results, the 2001 achievement-level results show higher percentages of eighth- and twelfth-grade students at or above both the *Basic* and *Proficient* levels among students whose parents graduated from college than among those who reported parents having lower levels of education.

Figure 3.10a

National Achievement Level Results by Parents' Education

Percentage of students within and at or above U.S. history achievement levels by parents' highest level of education, grade 8: 1994 and 2001



BEST COPY AVAILABLE

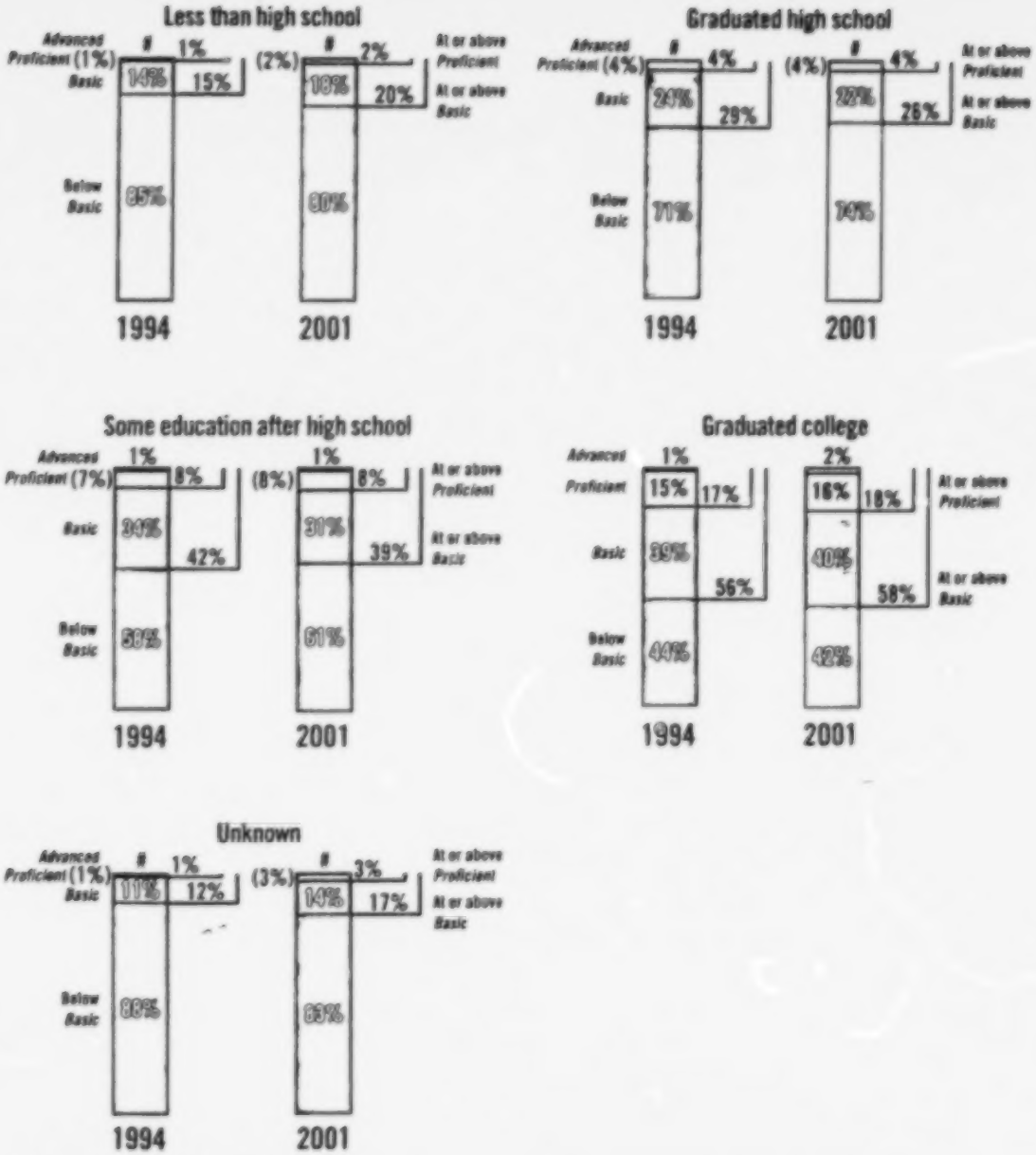
* Significantly different from 1994.
# Percentage is between 0.0 and 0.5.

NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

Figure 3.10b

National Achievement-Level Results by Parents' Education

Percentage of students within and at or above U.S. history achievement levels by parents' highest level of education, grade 12: 1994 and 2001



Percentage is between 0.0 and 0.5.

NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.

SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

Type of School

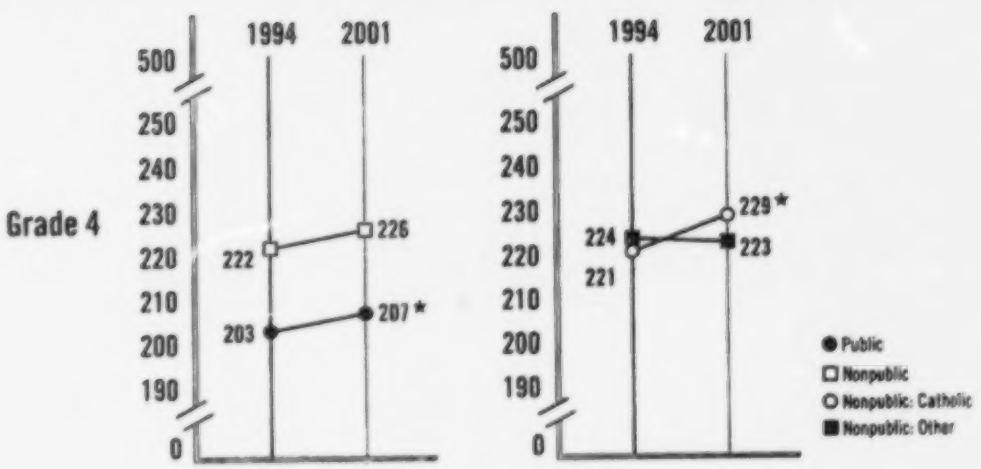
The schools that participate in the NAEP assessment are classified as either public or nonpublic. A further distinction is then made within the nonpublic classification between nonpublic schools that are Catholic and other nonpublic schools. In 2001, as in previous NAEP assessments, fourth-, eighth, and twelfth-grade students attending various types of nonpublic schools had higher average scores than did their peers attending public schools. Readers are, however, cautioned against making assumptions about the comparative quality of instruction in public and nonpublic schools. Socioeconomic and sociological factors that may affect student performance should be considered when interpreting these results.

Average U.S. history scale scores by type of school are presented in figure 3.11 and show higher average scores in 2001 than in 1994 among fourth- and eighth-grade students attending public schools. Fourth-graders attending Catholic nonpublic schools also had higher average scores in 2001 than in 1994. None of the apparent changes seen at grade 12 were statistically significant.

Comparisons of scale score results between the types of schools in 2001 show students at Catholic nonpublic schools outperforming public school students at all three grades, while the performance of students in other nonpublic schools was higher than that of public school students at grades 4 and 8 only.

Figure 3.11 — Average U.S. history scale scores by type of school, grades 4, 8, and 12: 1994 and 2001

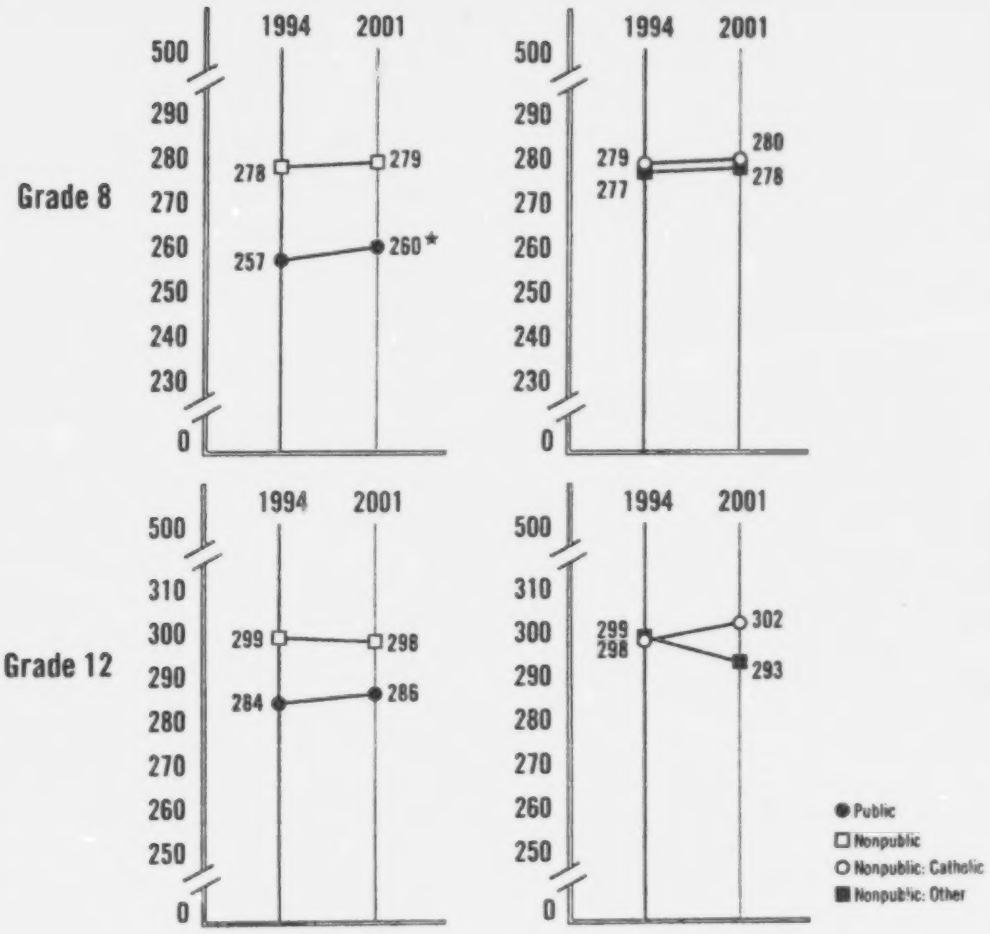
National Scale Score Results by Type of School



See footnotes at end of table. ▶

Figure 3.11 Average U.S. history scale scores by type of school, grades 4, 8, and 12: 1994 and 2001

National Scale Score Results by Type of School (continued)



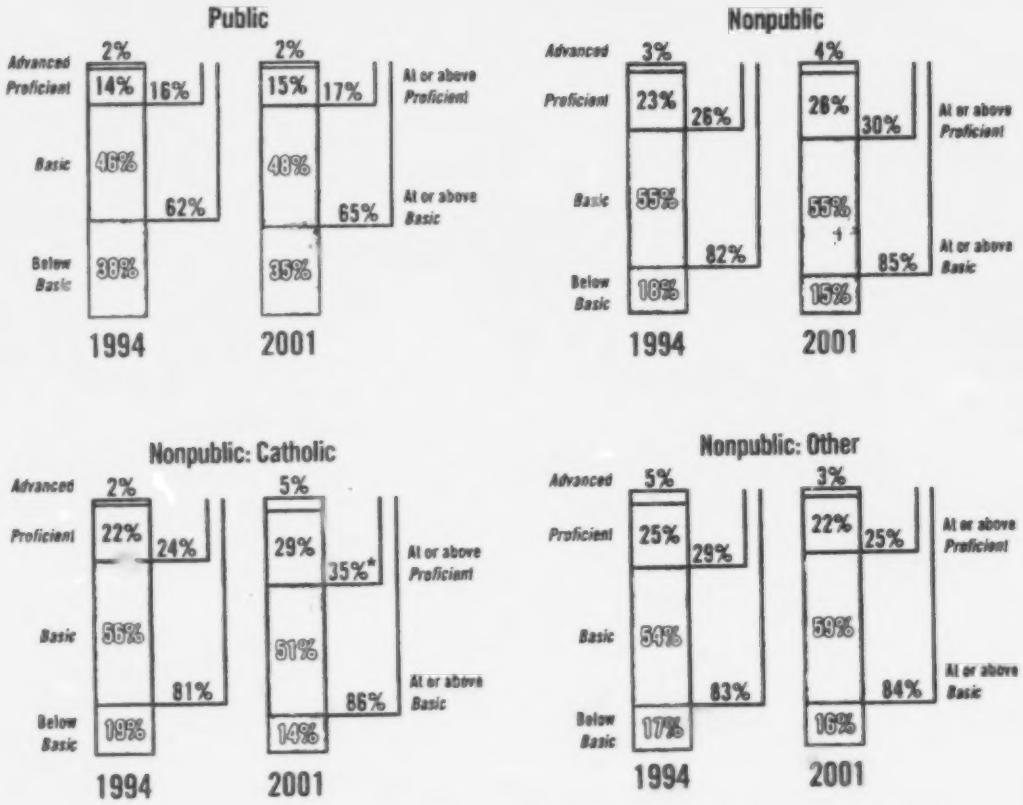
★ Significantly different from 1994.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

Achievement-level results by school type are presented in figure 3.12a, b, and c by grade. At grade 4, there was a higher percentage of Catholic nonpublic school students at or above *Proficient* in 2001 than in 1994. At grade 8, there were higher percentages of public school students at or above *Proficient* in 2001 than in 1994.

Comparisons of 2001 achievement-level results between types of schools show higher percentages of nonpublic school students at or above the *Basic* and *Proficient* levels than public school students at all three grades. There was also a higher percentage of eighth-grade students in nonpublic schools at the *Advanced* level than in public schools.

Figure 3.12a
National Achievement Level Results by Type of School

Percentage of students within and at or above U.S. history achievement levels by type of school, grade 4: 1994 and 2001

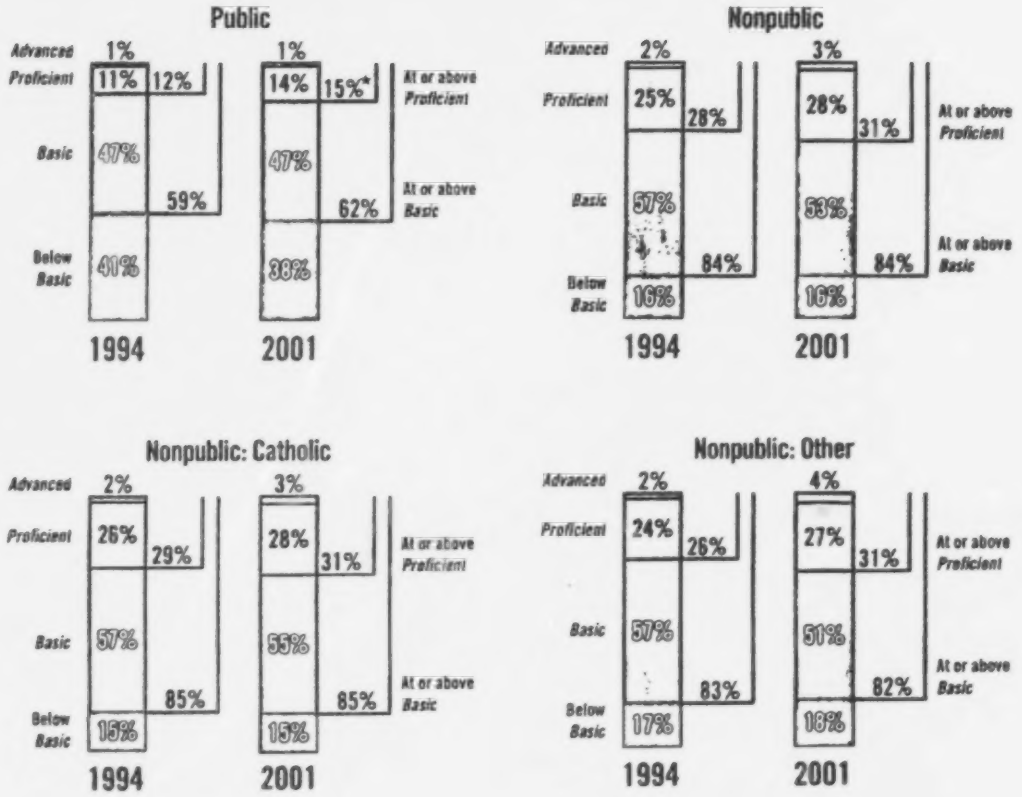


* Significantly different from 1994.
NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

Figure 3.12b

National Achievement-Level Results by Type of School

Percentage of students within and at or above U.S. history achievement levels by type of school, grade 8: 1994 and 2001



★ Significantly different from 1994.

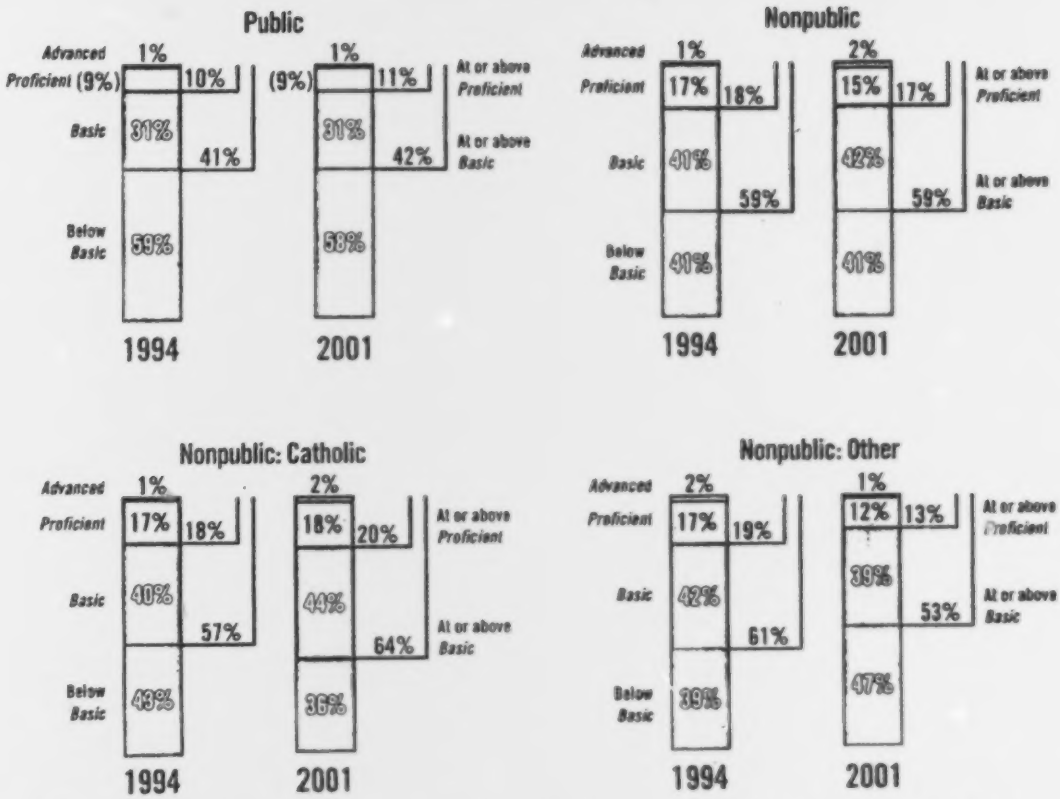
NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.

SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

Figure 3.12c

National Achievement-Level Results by Type of School

Percentage of students within and at or above U.S. history achievement levels by type of school, grade 12: 1994 and 2001



NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

BEST COPY AVAILABLE

Type of Location

The schools from which NAEP draws its samples of students are classified according to their type of location. Based on Census Bureau definitions of metropolitan statistical areas, including population size and density, the three mutually exclusive categories are: central city, rural/small town, and urban fringe/large town. Because of new methods used by NCES to identify the type of location assigned to each school in the Common Core of Data, schools were not classified in exactly the same way in 2001 as in 1994. Therefore, comparisons between the two assessment years are not possible, and only the data for the 2001 assessment are reported. More information on the definitions of the 2001 assessment classifications of location type is given in appendix A.

The performance of students by type of school location is shown in table 3.1. At grades 4 and 8, students attending schools in rural and urban fringe locations had higher average scores than students in central city schools. At grade 12, students attending schools in urban fringe locations had higher scores than students in both rural and central city locations.

Table 3.1 National Scale Score Results by Type of Location

Average U.S. history scale scores by type of location, grades 4, 8, and 12: 2001

	Central city	Urban fringe/large town	Rural/small town
Grade 4	199	211	215
Grade 8	257	265	263
Grade 12	283	292	284

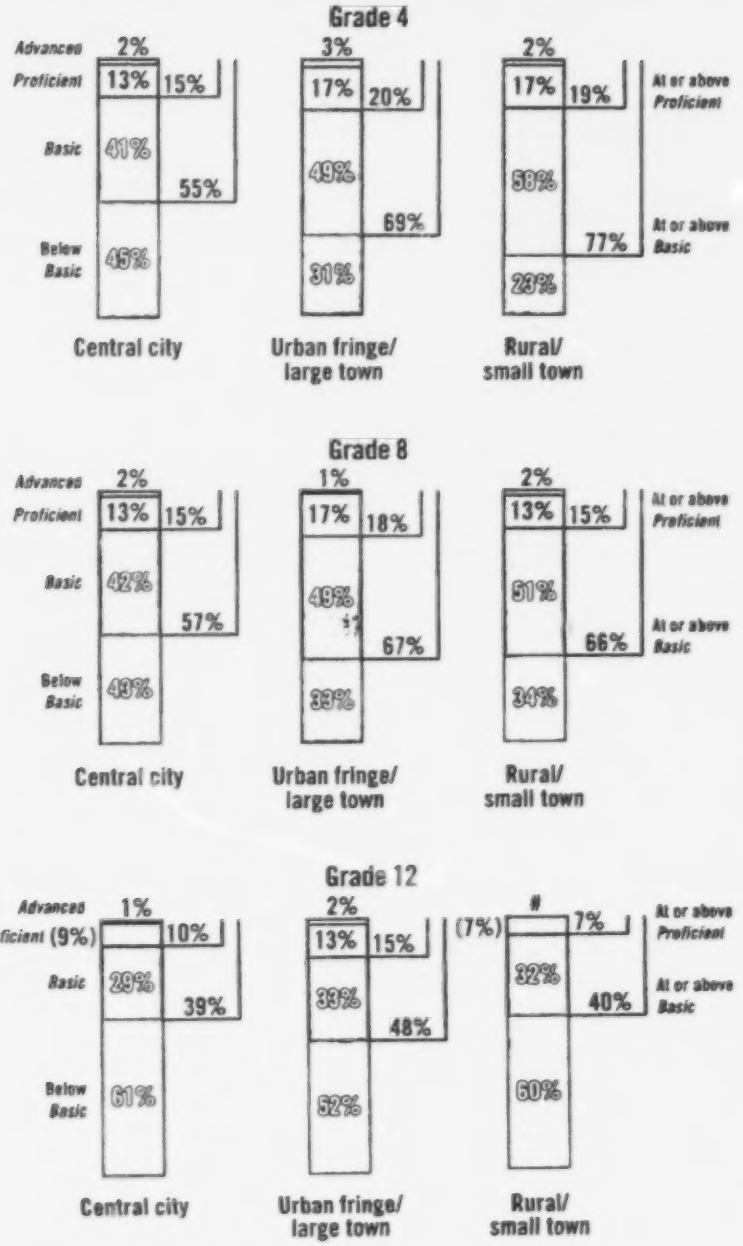
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Figure 3.13 presents achievement-level results by type of school location. Comparisons of achievement-level results between locations show higher percentages of fourth- and eighth-grade students at or above *Basic* in rural and urban fringe locations than in central city locations. There was also a higher percentage of fourth-graders at or above the *Basic* level in rural schools than in urban fringe schools. At grade 12, the percentages of students at or above the *Basic* and *Proficient* levels were higher in schools located in urban fringe areas than those in both rural and central city locations.

Figure 3.13

National Achievement-Level Results by Type of Location

Percentage of students within and at or above U.S. history achievement levels by type of school location, grades 4, 8, and 12: 2001



Percentage is between 0.0 and 0.5.

NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.

SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Free/Reduced-Price School Lunch Program Eligibility

Funded by the U.S. Department of Agriculture (USDA) as part of the National School Lunch Program, the Free/Reduced-Price School Lunch program is designed to assure that children at or near the poverty line receive nourishing meals. Eligibility guidelines for the lunch program are based on the Federal income poverty guidelines and are stated by household size.[1] NAEP first began collecting data on student eligibility for this program in 1996; therefore cross-year comparisons back to 1994 are not possible. Table 3.2 presents the 2001 U. S. history scale score results by students' eligibility for the program. At every grade, the average scale scores for students who were not eligible for the Free/Reduced-Price School Lunch program (i.e., those above the poverty guidelines) were higher than the scores for the students who are eligible for the program. Since information on eligibility is not available for a substantial percentage of the students at each grade, table 3.2 also displays the scale score averages for this third group of students. (Some schools do not offer free/reduced-price lunches. Students from these schools are counted in the Information Not Available category.) This group also had higher scale scores at every grade than the students eligible for the Free/Reduced-Price School Lunch program.

Table 3.2 National Scale Score Results by Free/Reduced-Price School Lunch Program Eligibility

Average U.S. history scale scores by student eligibility for Free/Reduced-Price School Lunch program, grades 4, 8, and 12: 2001

	Eligible	Not eligible	Info not available
Grade 4	189	220	217
Grade 8	245	269	268
Grade 12	271	289	295

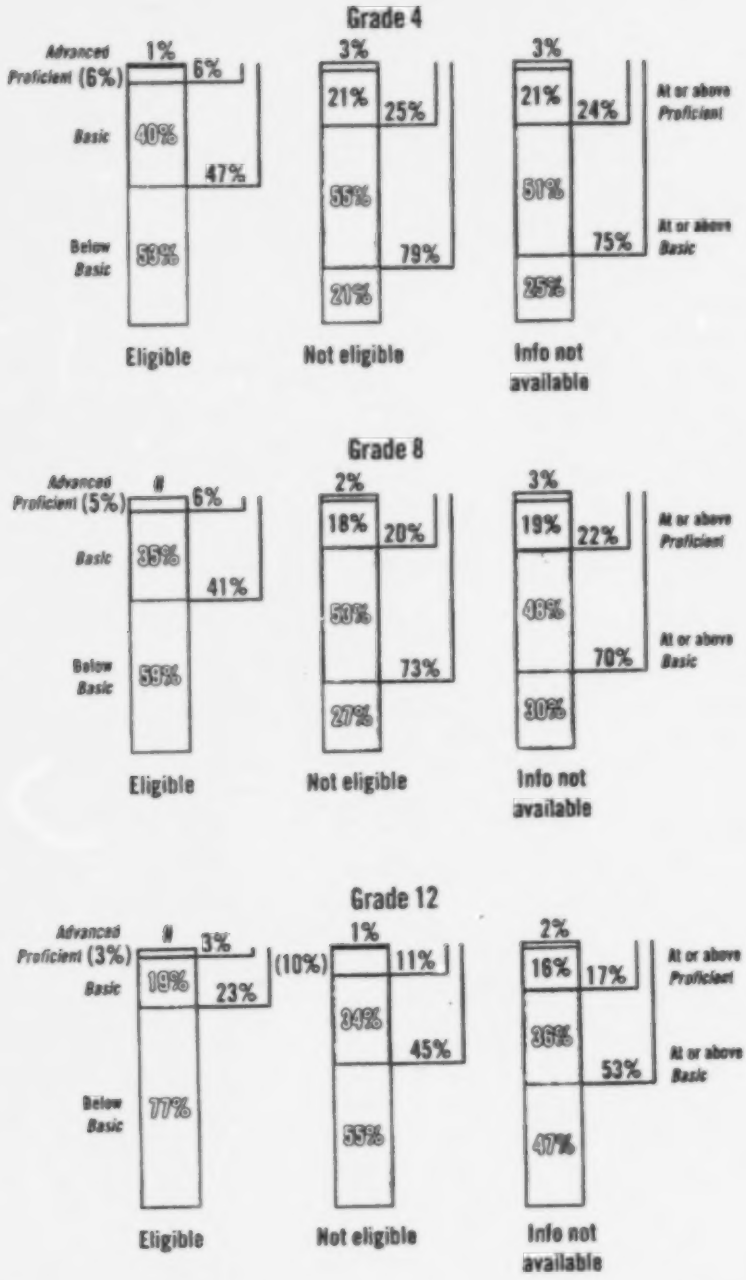
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

The pattern for achievement-level results is displayed in figure 3.14 and parallels that seen in the scale scores. There were higher percentages of fourth-, eighth, and twelfth-grade students at or above *Basic* and *Proficient* who were not eligible for the program than those who were eligible for it.

[1] U.S. General Services Administration. (1999). *Catalogue of federal domestic assistance.* Washington, DC: Executive Office of the President, Office of Management and Budget.

Figure 3.14
National Achievement-Level Results by Free/Reduced-Price School Lunch Program Eligibility

Percentage of students within and at or above U.S. history achievement levels by Free/Reduced-Price School Lunch program eligibility, grades 4, 8, and 12: 2001



Percentage is between 0.0 and 0.5.
NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

BEST COPY AVAILABLE

4 Becoming a More Inclusive National Assessment

In its efforts to assess a representative sample of all students in the nation, NAEP consistently has striven to include special-needs students—those with disabilities (SD) or limited English proficient students (LEP). A certain percentage of such students, however, has always been excluded because they could not be assessed meaningfully without accommodations. Schools that participate in NAEP have been asked to use specific criteria in making decisions to exclude certain students who have been classified as having a disability under the Individuals with Disabilities Education Act (IDEA), based upon their Individualized Education Programs (IEP) and Section 504 of the Rehabilitation Act of 1973. Similarly, schools have been permitted to exclude some students they identify as being limited English proficient.[1]

Chapter Focus

How would the NAEP results differ if accommodations were permitted for special-needs students?

In order to increase the inclusiveness of NAEP's samples, and in an attempt to remain consistent with state- and district-level testing policies that increasingly offer accommodations to special-needs students, NAEP began to explore the use of accommodations in the 1996 and 1998 assessments. A split-sample design was used to identify a portion of schools that were permitted to provide accommodations to their special-needs students who required them, and a portion of schools in which accommodations were not offered (the

Chapter Contents

Two Sets of 2001 NAEP U.S. History Results

Results for the Nation

National Results by Gender

National Results by Race/Ethnicity

[1] See appendix A for a description of specific criteria provided to assist them in making exclusion decisions.

BEST COPY AVAILABLE

standard administration procedure prior to 1996). The split-sample design made it possible to study the effects on NAEP results of including special-needs students who required and were provided accommodations, while at the same time, obtaining results that were comparable to those from previous assessments. Based on research conducted and published since that time, it was determined that NAEP could begin a transition to reporting results that included the performance of accommodated special-needs students.[2] It is anticipated that in the near future, NAEP will only report results based on this more inclusive sample.

Two Sets of 2001 NAEP U.S. History Results

This report is the first to display two different sets of NAEP U.S. history results based on the split-sample design: 1) those that reflect the performance of regular and special-needs students when accommodations were not permitted, and 2) those that reflect the performance of regular and special-needs students—both those who were accommodated and those who could be tested without accommodations—when accommodations were permitted. It should be noted that accommodated students make up a small proportion of the total weighted number of students assessed (see table A.6, page 115 in appendix A for details). Making accommodations available may change the overall assessment results in subtle and different ways. For example, when accommodations are permitted, there may be some occurrences of students being accommodated who might have taken the test under standard conditions if accommodations were not permitted. This could lead to an overall increase in the average assessment results if accommodations were to increase special-needs students' performance. Conversely, when accommodations are permitted, special-needs students who could not have been tested without accommodations could be included in the sample. Assuming that these are generally lower-performing students, their inclusion in the sample—even with accommodations—could result in an overall lower average score.

The two sets of results presented in this chapter were obtained by administering the assessment to a nationally representative sample of students and schools. In one sample, no accommodations were permitted; all students were assessed under the same conditions that were the basis for reporting results from the 1994 NAEP U.S. history assessment. In another part of the

---

2 Olson, J. F. & Goldstein, A. A. (1997). *The inclusion of students with disabilities and limited-English-proficient students in large-scale assessments: A summary of recent progress.* (NCES Publication No. 97–482). Washington, DC: National Center for Education Statistics.

Mazzeo, J., Carlson, J. E., Voelkl, K. E., & Lutkus, A. D. (1999). *Increasing the participation of special needs students in NAEP: A report on 1996 research activities.* (NCES Publication No. 2000–473). Washington, DC: National Center for Education Statistics.

schools sampled, accommodations were permitted for students with disabilities and limited English proficient students who normally receive accommodations in their district or state assessment programs. Most accommodations that schools routinely provide for their own testing programs were permitted. The permitted accommodations included, but were not limited to the following:

- one-on-one testing,
- bilingual dictionary,
- large print book,
- small-group testing,
- extended time,
- oral reading of questions, and
- use of an aide for transcribing responses.

(See appendix A, table A.7, page 117, for greater detail on the numbers and percentages of students accommodated by accommodation type in the 2001 assessment.)

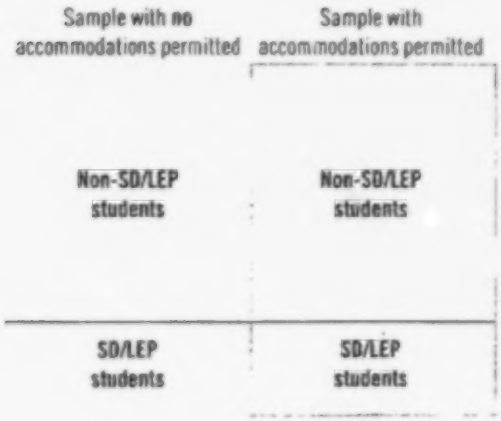
Figure 4.1 provides a visual representation of how the two sets of results were based on the two samples in 2001. Included in both sets of results (accommodations not permitted and accommodations permitted) are those students from both samples of schools who were not identified as either SD or LEP. In addition, the first set of results (accommodations not permitted) includes SD and LEP students from the sample of schools where accommodations were not permitted (see middle portion of figure 4.1). This is the set of results that allows for trend comparisons back to 1994 and are presented in the other chapters of this report.

The second set of results, accommodations permitted (see bottom portion of figure 4.1), includes SD and LEP students from the sample of schools where accommodations were permitted. This is the set of results that form the new, more inclusive baseline for future reporting of trend comparisons for the NAEP U.S. history assessment.

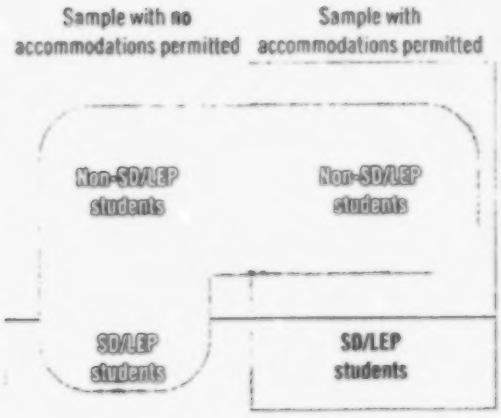
In the NAEP 2001 sample where accommodations were not permitted, 16 percent of fourth-graders, 16 percent of eighth-graders, and 11 percent of twelfth-graders, were identified by their schools as having special needs (i.e., either as students with disabilities or limited English proficient students). In the other sample where accommodations were offered, 18 percent of fourth-graders, 17 percent of eighth-graders, and 10 percent of twelfth-graders were identified as having special needs. In the sample where accommodations were not permitted, between 45 and 51 percent of the special-needs students at each of the three grade levels (between 4 and 8 percent of all students—see appendix A, table A.5, page 114) were excluded from NAEP testing by their schools. In the sample where accommodations were offered, between 23 and 33 percent of the special-needs students were excluded from the assessment (between 2 and 3 percent of the

Figure 4.1 Split-Sample Design

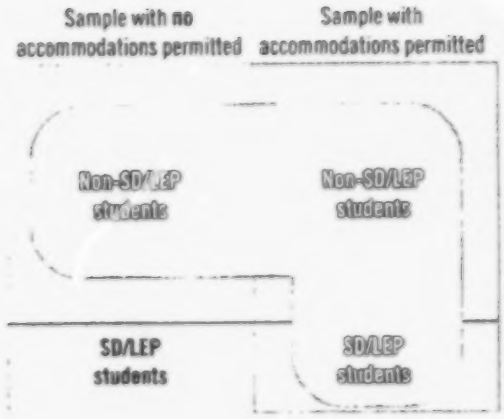
The two sets of NAEP results based on a split-sample design



Split-sample design
The national sample was split. In part of the schools, accommodations were not permitted for students with disabilities (SD) and limited English proficient (LEP) students. In the other schools, accommodations were permitted for SD and LEP students who routinely received them in their school assessments.



Accommodations-not-permitted results
The accommodations-not-permitted results include the performance of students from both samples who were not classified as SD or LEP and the performance of SD and LEP students from the sample in which no accommodations were permitted.



Accommodations-permitted results
The accommodations-permitted results also include the performance of students from both samples who were not classified as SD or LEP; however, the SD and LEP students whose performance is included in this set of results were from the sample in which accommodations were permitted. Since students who required testing accommodations could be assessed and represented in the overall results, it was anticipated that these results would include more special-needs students and reflect a more inclusive sample.

BEST COPY AVAILABLE

total sample). Thus, offering accommodations would appear to lead to greater inclusion of special-needs students.

Chapters 2, 3, 5, and 6 of this report are based on the first set of results (no accommodations permitted). This chapter presents an overview of the second set of results—results that include students who were provided accommodations during the assessment administration. Overall results are provided for the nation and for student subgroups by gender and by race/ethnicity. These results are discussed in terms of statistically significant differences between the two sets of results and differences between subgroups of students within each set of results. Throughout this chapter, the assessment results that include SD and LEP students for whom accommodations were not permitted will be referred to as the "accommodations-not-permitted" results. The set of results that includes SD and LEP students for whom accommodations were permitted will be referred to as the "accommodations-permitted" results.

Results for the Nation

Accommodations Not Permitted and Accommodations Permitted

Table 4.1 displays the average U.S. history scale scores for the nation in 2001 for two sets of results: 1) accommodations not permitted, and 2) accommodations permitted. There were no significant differences in the average scores between the two sets of results at grades 4 and 12. At grade 8, however, the average score when accommodations were permitted was lower than the average score when accommodations were not permitted.

As noted in the introduction to this chapter, NAEP has always sought to include special-needs students proportional to their representation in the U.S. population. Offering accommodations tends to reduce exclusion rates for special-needs students and therefore allows NAEP to offer a fairer and more accurate picture of the status of American education. Because special-needs students are typically classified as eligible for special educational

Table 4.1 Comparison of Two Sets of National Scale Score Results

National average U.S. history scale scores by type of results, grades 4, 8, and 12: 2001

	Accommodations not permitted	Accommodations permitted
Grade 4	209	208
Grade 8	262	260 †
Grade 12	287	287

† Significantly different from the sample where accommodations were not permitted.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

services after having shown some difficulty in the regular learning environment, some may assume that including the performance of these students would tend to lower the overall results. This assumption appears to have been justified only in the observed difference between the two sets of grade 8 U.S. history results in 2001, where the accommodations-permitted results, which included slightly more special-needs students because of the availability of accommodations, were lower than the accommodations-not-permitted results. It is important to examine the percentages of students attaining the NAEP achievement levels, however, to see if there were higher percentages at the lower achievement level (i.e., *Basic*), when students were assessed with accommodations.

Table 4.2 shows the percentages of students attaining each of the achievement levels. The percentages are similar across the two sets of results for grades 4 and 12; apparent differences between the accommodations-not-permitted and the accommodations-permitted results were not significantly different. At grade 8, however, the percentage of students below *Basic* was higher when accommodations were permitted than when they were not permitted.

BEST COPY AVAILABLE

Table 4.2 Comparison of Two Sets of National Achievement-Level Results

Percentage of students within and at or above U.S. history achievement levels by type of results, grades 4, 8, and 12: 2001

	Below *Basic*	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
Grade 4						
Accommodations were not permitted	33	49	16	2	67	18
Accommodations were permitted	34	48	16	2	66	18
Grade 8						
Accommodations were not permitted	36	48	15	2	64	17
Accommodations were permitted	38 †	46 †	14	1	62 †	16
Grade 12						
Accommodations were not permitted	57	32	10	1	43	11
Accommodations were permitted	57	32	10	1	43	11

† Significantly different from the sample where accommodations were not permitted.
NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

National Results by Gender

Accommodations Not Permitted and Accommodations Permitted

The average U.S. history scale scores by gender for both sets of results in 2001 are provided in table 4.3. Both male and female students at grade 8 had higher U.S. history scores when accommodations were not permitted than when accommodations were permitted.

At all three grades, the average scores for male students were not significantly different from that of female students regardless of whether or not accommodations were permitted.

Table 4.3 Comparison of Two Sets of National Scale Score Results by Gender

National average U.S. history scale scores by gender and type of results, grades 4, 8, and 12: 2001

	Male	Female
Grade 4		
Accommodations were not permitted	209	209
Accommodations were permitted	207	209
Grade 8		
Accommodations were not permitted	264	261
Accommodations were permitted	261 †	260 †
Grade 12		
Accommodations were not permitted	288	286
Accommodations were permitted	288	286

† Significantly different from the sample where accommodations were not permitted.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

BEST COPY AVAILABLE

The percentages of male and female students attaining the *Basic*, *Proficient*, and *Advanced* levels are provided in table 4.4. Comparing the two sets of results in 2001, there were no statistically significant differences in the percentages of male or female students attaining each of the achievement levels at grades 4 or 12. At grade 8, however, a higher percentage of male students were below *Basic* when accommodations were permitted than when they were not.

Table 4.4 Comparison of Two Sets of National Achievement-Level Results by Gender

Percentage of students within and at or above U.S. history achievement levels by gender and type of results, grades 4, 8, and 12: 2001

	Below *Basic*	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
Grade 4						
Male						
Accommodations were not permitted	34	47	17	2	66	19
Accommodations were permitted	35	46	16	2	65	19
Female						
Accommodations were not permitted	32	51	15	2	68	17
Accommodations were permitted	33	50	15	2	67	17
Grade 8						
Male						
Accommodations were not permitted	35	47	17	2	65	18
Accommodations were permitted	38 †	45	16	2	62 †	17
Female						
Accommodations were not permitted	37	48	14	1	63	15
Accommodations were permitted	39	47	13	1	61	14
Grade 12						
Male						
Accommodations were not permitted	55	33	11	1	45	12
Accommodations were permitted	55	32	11	2	45	12
Female						
Accommodations were not permitted	59	31	9	1	41	10
Accommodations were permitted	60	31	9	1	40	10

† Significantly different from the sample where accommodations were not permitted.
NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

BEST COPY AVAILABLE

National Results by Race/Ethnicity

Accommodations Not Permitted and Accommodations Permitted

NAEP assessments across academic subjects have typically reported large score differences according to race and ethnic group membership. If students with disabilities or limited English proficient students are over-represented in a particular racial or ethnic group, that group's assessment scores may decrease. Table 4.5 provides the average U. S. history scale scores for each of the race/ethnicity categories for the two sets of results in 2001. At grade 8, both White students and Black students had higher average scores when accommodations were not permitted than when accommodations were permitted. There were no statistically significant differences observed between the average scores when accommodations were not permitted and when accommodations were permitted for any of the race/ethnicity categories at grades 4 and 12.

As noted in chapter 3, a pattern of performance differences by race/ethnicity can be seen in the accommodations-not-permitted results in 2001. Both White and Asian/Pacific Islander students at all three grades scored higher than Black and Hispanic students. The same pattern can be observed in the accommodations-permitted results. However, while White students outperformed their Asian/Pacific Islander peers at grade 4 when accommodations were not permitted, the difference was not statistically significant when accommodations were permitted.

Table 4.5 Comparison of Two Sets of National Scale Score Results by Race/Ethnicity

National average U.S. history scale scores by race/ethnicity and type of results, grades 4, 8, and 12: 2001

	White	Black	Hispanic	Asian/Pacific Islander	American Indian
Grade 4					
Accommodations were not permitted	220	188	186	213	197
Accommodations were permitted	218	186	187	214	197
Grade 8					
Accommodations were not permitted	271	243	243	267	249
Accommodations were permitted	269 †	240 †	240	265	248
Grade 12					
Accommodations were not permitted	292	269	274	295	277
Accommodations were permitted	292	268	271	294	274

† Significantly different from the sample where accommodations were not permitted.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

The percentages of students in each race/ethnicity category who attained the *Basic*, *Proficient*, and *Advanced* levels are provided in table 4.6. No significant differences were found at any of the three grades between the accommodations-not-permitted results and the accommodations-permitted results for the percentages of students attaining each of the achievement levels in 2001.

BEST COPY AVAILABLE

Table 4.6 Comparison of Two Sets of National Achievement-Level Results by Race/Ethnicity

Percentage of students within and at or above U.S. history achievement levels by race/ethnicity and type of results, grades 4, 8, and 12: 2001

	Below *Basic*	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
Grade 4						
White						
Accommodations were not permitted	21	55	21	3	79	24
Accommodations were permitted	23	53	21	3	77	24
Black						
Accommodations were not permitted	56	38	5	#	44	6
Accommodations were permitted	58	36	5	#	42	5
Hispanic						
Accommodations were not permitted	58	35	6	1	42	7
Accommodations were permitted	58	36	6	#	42	6
Asian/Pacific Islander						
Accommodations were not permitted	29	53	16	3	71	19
Accommodations were permitted	26	54	17	3	74	20
American Indian						
Accommodations were not permitted	47	41	8	4	53	12
Accommodations were permitted	44	44	9	3	56	12
Grade 8						
White						
Accommodations were not permitted	25	53	19	2	75	21
Accommodations were permitted	27	52	19	2	73	20
Black						
Accommodations were not permitted	62	34	4	#	38	4
Accommodations were permitted	65	31	4	#	35	4
Hispanic						
Accommodations were not permitted	60	34	5	#	40	5
Accommodations were permitted	63	32	4	#	37	4
Asian/Pacific Islander						
Accommodations were not permitted	32	48	18	2	68	20
Accommodations were permitted	34	47	17	2	66	19
American Indian						
Accommodations were not permitted	50	42	7	1	50	8
Accommodations were permitted	54	38	7	1	46	8
Grade 12						
White						
Accommodations were not permitted	51	36	12	1	49	13
Accommodations were permitted	51	36	12	1	49	13
Black						
Accommodations were not permitted	80	18	3	#	20	3
Accommodations were permitted	80	17	3	#	20	3
Hispanic						
Accommodations were not permitted	74	21	5	#	26	5
Accommodations were permitted	74	21	5	#	26	5
Asian/Pacific Islander						
Accommodations were not permitted	47	31	17	5	53	21
Accommodations were permitted	48	31	16	5	52	21
American Indian						
Accommodations were not permitted	66	33	1	0	34	1
Accommodations were permitted	68	31	1	0	32	1

Percentage is between 0.0 and 0.5.

NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.

SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

BEST COPY AVAILABLE

5 Sample Assessment Questions and Student Responses

This chapter presents sample questions from the 2001 NAEP U.S. history assessment and examples of student responses to those questions. Four sample questions at each grade level are provided, including multiple-choice, short constructed-response, and extended constructed-response questions. Each sample question is classified according to a historical theme or historical period, as described in the U.S. history framework. Actual student responses to the sample constructed-response questions have been reproduced from test booklets to illustrate answers representative of the indicated scoring categories.

The tables accompanying each sample question present two types of performance data: the overall percentage of students who answered successfully and the percentage of students within a specific score range on the NAEP U.S. history scale who answered successfully. The score ranges presented are those that correspond to the three achievement-level intervals—*Basic, Proficient,* and *Advanced*—as well as the score range that falls below *Basic.*

Chapter Focus

Materials from the 2001 U.S. history assessment

Chapter Contents

Sample Questions

Student Responses

Item Maps

BEST COPY AVAILABLE

Grade 4 Sample Assessment Question Results

Assessment questions at grade 4, as at the other grade levels, included both constructed-response and multiple-choice formats. Many, like the examples below, used visual or textual stimuli. Questions tended to concern topics that are typically addressed in fourth-grade social studies classes.

Grade 4 Sample Question 1:

In pioneer schools, feathers like this were most often used for

- (A) measuring
- (B) sewing
- (C) writing
- (D) playing a game

Historical Theme:	Historical Period:
Economic and Technological Changes and Their Relation to Society, Ideas, and the Environment	Expansion and Reform (1801 to 1861)

Table 5.1 Sample Question 1 Results (Multiple-Choice)

Overall percentage correct and percentages correct within each achievement level range: 2001

Grade 4	Percentage correct within achievement-level intervals			
Overall percentage correct	Below *Basic* 194 and below*	*Basic* 195–242*	*Proficient* 243–275*	*Advanced* 276 and above*
93	84	96	99	***

*NAEP U.S. history composite scale range.
***Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

BEST COPY AVAILABLE

Grade 4 Sample Question 2:

The poster shown above is trying to attract recruits by appealing to their

- A homesickness
- B religious beliefs
- C patriotism
- D need for money

Historical Theme:	Historical Period:
The Changing Role of America in the World	Modern America and the World Wars (1914 to 1945)

Table 5.2 Sample Question 2 Results (Multiple-Choice)

Overall percentage correct and percentages correct within each achievement level range: 2001

Grade 4	Percentage correct within achievement-level intervals			
Overall percentage correct	Below *Basic* 194 and below*	*Basic* 195–242*	*Proficient* 243–275*	*Advanced* 276 and above*
45	25	46	74	***

*NAEP U.S. history composite scale range.
***Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

BEST COPY AVAILABLE

Grade 4 Sample Question 3:

What was a major cause of the Civil War?

- (A) People in the North and in the South had different religions.
- (B) People in the North and in the South disagreed over slavery.
- (C) People in the North wanted control of the country when they found out that gold had been discovered in the South.
- (D) People in the South wanted control of the country when they found out that oil had been discovered in the North.

Historical Theme:	Historical Period:
Change and Continuity in American Democracy: Ideas, Institutions, Practices, and Controversies	Crisis of the Union: Civil War and Reconstruction (1850 to 1877)

Table 5.3 Sample Question 3 Results (Multiple-Choice)

Overall percentage correct and percentages correct within each achievement level range: 2001

Grade 4	Percentage correct within achievement-level intervals			
Overall percentage correct	Below *Basic* 194 and below*	*Basic* 195–242*	*Proficient* 243–275*	*Advanced* 276 and above*
57	34	59	86	***

*NAEP U.S. history composite scale range.
***Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

BEST COPY AVAILABLE

Grade 4 Sample Question 4

Choose an American Indian group from the map, and circle its name directly on the map.

On the chart below, list one way this American Indian group got food, shelter, and clothing in the period before Europeans came to the Americas. Then list one way your family gets food, shelter, and clothing.

	American Indians in the Period Before Europeans Came	Your Family
1.	Food: _____	Food: _____
2.	Shelter: _____	Shelter: _____
3.	Clothing: _____	Clothing: _____

Give one reason why the American Indian group long ago and your family today differ in the ways they get their food, shelter, or clothing.

_____

_____

Historical Theme	Historical Period
The Gathering and Interactions of Peoples, Cultures, and Ideas	Three Worlds and Their Meeting in the Americas (Beginnings to 1607)

BEST COPY AVAILABLE

Responses to this question were scored according to a four-level rubric as Inappropriate, Partial, Essential, or Complete.

Table 5.4 Sample Question 4 Results (Extended Constructed-Response)

Overall percentage "Essential" or better and percentages "Essential" or better within each achievement level range: 2001

Grade 4	Percentage "Essential" or better within achievement-level intervals			
Overall percentage "Essential" or better	Below *Basic* 194 and below*	*Basic* 195–242*	*Proficient* 243–275*	*Advanced* 276 and above*
42	13	48	76	***

*NAEP U.S. history composite scale range.
***Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

BEST COPY AVAILABLE

Responses scored "Complete" circled an American Indian group and correctly listed ways that the group chosen got food, shelter, and clothing, and ways that the student's own family gets food, shelter, and clothing. They then gave one appropriate reason for differences between the way the American Indian group obtained those necessities and the way in which modern families obtain them.

Sample "Complete" Response:

American Indians in the Period Before Europeans Came		Your Family	
1. Food:	They planted food.	1. Food:	We go to the supermarket.
2. Shelter:	They built houses made of trees	2. Shelter:	We buy finished built houses.
3. Clothing:	They used the skin of the animals they killed.	3. Clothing:	We buy from shopping malls.

Give one reason why the American Indian group long ago and your family today differ in the ways they get their food, shelter, or clothing.

We differ because a long time ago there weren't any stores or real estate people as we have today.

BEST COPY AVAILABLE

Responses scored "Essential" circled an American Indian group and correctly listed two ways that the group chosen got food, shelter, or clothing, and two ways that modern families obtain them. They did not give an appropriate reason for these differences.

Sample "Essential" Response:

American Indians in the Period Before Europeans Came

1. Food: They Hunted
2. Shelter: They used Buffalo hide
3. Clothing: Animal furs

Your Family

1. Food: Grocery Store
2. Shelter: Houses
3. Clothing: Department Stores

Give one reason why the American Indian group long ago and your family today differ in the ways they get their food, shelter, or clothing.

We can buy things today.

BEST COPY AVAILABLE

Responses scored "Partial" circled an American Indian group and were able to identify correctly one way the group chosen got food, shelter, or clothing and one way modern families get them.

Sample "Partial" Response:

American Indians in the Period Before Europeans Came		Your Family	
1. Food:	They grew crops.	1. Food:	We buy it.
2. Shelter:	They made it!	2. Shelter:	We buy it too
3. Clothing:		3. Clothing:	We could make it.

Give one reason why the American Indian group long ago and your family today differ in the ways they get their food, shelter, or clothing.

They might make it and do what there left to do with it.

BEST COPY AVAILABLE

Grade 8 Sample Assessment Question Results

In addition to assessing content appropriate to an eighth-grade U.S. history course, assessment questions for eighth grade assessed a range of history skills, such as the text interpretation skills measured in sample question 8.

Grade 8 Sample Question 5:

Why was Roger Williams forced to leave the Massachusetts Bay Colony?

- ○ He claimed that the Puritan government had no right to control religious beliefs.
- ⓑ He was more loyal to the King of Spain than to the English monarchy.
- ⓒ He refused to do his share of the farming and other work.
- ⓓ He wanted to lead a war against the American Indians.

Historical Theme:	Historical Period:
Change and Continuity in American Democracy: Ideas, Institutions, Practices, and Controversies	Colonization, Settlement, and Communities (1607 to 1763)

Table 5.5 Sample Question 5 Results (Multiple-Choice)

Overall percentage correct and percentages correct within each achievement level range: 2001

Grade 8	Percentage correct within achievement-level intervals			
Overall percentage correct	Below *Basic* 251 and below*	*Basic* 252–293*	*Proficient* 294–326*	*Advanced* 327 and above*
52	34	56	79	***

*NAEP U.S. history composite scale range.
***Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

BEST COPY AVAILABLE

Grade 8 Sample Question 6:

What was the most significant factor that led the American colonists to form the First Continental Congress in 1774 ?

- (A) Religious conflict inside the colonies
- (B) The desire of the colonists to write a Constitution to replace the Articles of Confederation
- (C) Colonial frustration with laws passed by the British Parliament
- (D) The desire of the colonists to stop the war between Britain and the colonies

Historical Theme:	Historical Period:
Change and Continuity in American Democracy: Ideas, Institutions, Practices, and Controversies	The Revolution and the New Nation (1763 to 1815)

Table 5.6 Sample Question 6 Results (Multiple-Choice)

Overall percentage correct and percentages correct within each achievement level range: 2001

Grade 8	Percentage correct within achievement-level intervals			
Overall percentage correct	Below *Basic* 251 and below*	*Basic* 252–293*	*Proficient* 294–326*	*Advanced* 327 and above*
39	29	39	62	***

*NAEP U.S. history composite scale range.
***Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

BEST COPY AVAILABLE

Grade 8 Sample Question 7:

Why was the invention of the steel plow important in United States history?

_____

_____

_____

Historical Theme:	Historical Period:
Economic and Technological Changes and Their Relation to Society, Ideas, and the Environment	The Development of Modern America (1805 to 1920)

Responses to this question were scored according to a three-level rubric as Inappropriate, Partial, or Appropriate

Table 5.7 Sample Question 7 Results (Short Constructed-Response)

Overall percentage "Appropriate" and percentages "Appropriate" within each achievement level range: 2001

Grade 8	Percentage "Appropriate" within achievement-level intervals			
Overall percentage "Appropriate"	Below *Basic* 251 and below*	*Basic* 252–293*	*Proficient* 294–326*	*Advanced* 327 and above*
30	9	34	64	***

*NAEP U.S. history composite scale range.
***Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

This "Appropriate" response indicated that the steel plow increased efficiency in agricultural production.

Sample "Appropriate" Response:

Why was the invention of the steel plow important in United States history?

The steel plow was stronger, lasted longer, worked faster, and could farm harder ground.

This "Partial" response indicated in a general way that the steel plow made farming easier; it was correct but was not specific about the steel plow's impact.

Sample "Partial" Response:

Why was the invention of the steel plow important in United States history?

to make faerming a lot easier.

Grade 8 Sample Question 8:

> In America, it was no disgrace to work at a trade. Workmen and capitalists were equal. The employer addressed the employee as you, not familiarly as thou. The cobbler and the teacher had the same title, "mister," and all the children, boys and girls, Jews and Gentiles, went to school!
>
> —Polish immigrant, 1910

Using the quotation above and your knowledge of history, explain in your own words two important aspects of life in the United States that seemed good to this immigrant.

1) _____

_____

2) _____

_____

What do you think was the most important difference this man saw between Poland and the United States?

_____

_____

_____

Historical Theme:	Historical Period:
The Gathering and Interactions of Peoples, Cultures, and Ideas	The Development of Modern America (1865 to 1920)

Responses to this question were scored according to a three-level rubric as Inappropriate, Partial, or Appropriate

Table 5.8 Sample Question 8 Results (Short Constructed-Response)

Overall percentage "Appropriate" and percentages "Appropriate" within each achievement level range: 2001

Grade 8	Percentage "Appropriate" within achievement-level intervals			
Overall percentage "Appropriate"	Below *Basic* 251 and below*	*Basic* 252–293*	*Proficient* 294–326*	*Advanced* 327 and above*
33	11	38	64	***

*NAEP U.S. history composite scale range.
***Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Responses scored "Appropriate" correctly identified two positive aspects of American life as perceived by the immigrant quoted and explained an important distinction between the United States and Poland in the eyes of that immigrant.

Sample "Appropriate" Response:

Using the quotation above and your knowledge of history, explain in your own words two important aspects of life in the United States that seemed good to this immigrant.

1) The Polish immigrant felt that no one was shut out because of Religion

2) And no matter what job you had you were treated fairly

What do you think was the most important difference this man saw between Poland and the United States?

The most important difference is how the employer sees the employee

BEST COPY AVAILABLE

Responses scored "Partial" correctly identified two positive aspects of American life as perceived by the immigrant quoted but did not explain an important distinction between the United States and Poland in the eyes of the immigrant.

Sample "Partial" Response:

Using the quotation above and your knowledge of history, explain in your own words two important aspects of life in the United States that seemed good to this immigrant.

1) Everyone was alike

2) Everyone was live the same respect.

What do you think was the most important difference this man saw between Poland and the United States?

Some were from Poland some were from the U.S.

BEST COPY AVAILABLE

Grade 12 Sample Assessment Question Results

Questions at the twelfth-grade level assessed U.S. history knowledge and skills at a more sophisticated level than those at grades 4 and 8. This greater complexity is presented in the questions included below both in the degree of detail required and the extent to which students must understand and analyze historical issues from various perspectives, as is evident in the constructed-response questions.

Grade 12 Sample Question 9

The Progressive movement of 1890-1920 is best described as

- (A) a broad-based reform movement that tried to reduce the abuses that had come with modernization and industrialization
- (B) a loose coalition of groups primarily dedicated to passing a constitutional amendment prohibiting the consumption of alcohol
- (C) an anti-tariff movement led by a federation of business owners and manufacturers who wanted to promote trade abroad
- (D) a grass-roots movement that attempted to gather support for the establishment of an international organization such as the League of Nations

Historical Theme:	Historical Period:
Change and Continuity in American Democracy: Ideas, Institutions, Practices, and Controversies	The Development of Modern America (1865 to 1920)

Table 5.9 Sample Question 9 Results (Multiple-Choice)

Overall percentage correct and percentages correct within each achievement level range: 2001

Grade 12	Percentage correct within achievement-level intervals			
Overall percentage correct	Below *Basic* 293 and below*	*Basic* 294–324*	*Proficient* 325–354*	*Advanced* 355 and above*
36	23	47	73	***

*NAEP U.S. history composite scale range.
***Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

BEST COPY AVAILABLE

Grade 12 Sample Question 10:

The phrase "Harlem Renaissance" refers to

- (A) African American political gains during the Reconstruction period
- (B) African American achievements in art, literature, and music in the 1920's
- (C) a religious revival in the African American community that swept the nation in the 1950's
- (D) a series of urban renewal projects that were part of the Great Society program of the 1960's

Historical Theme:	Historical Period:
The Gathering and Interactions of Peoples, Cultures, and Ideas	Modern America and the World Wars (1914 to 1945)

Table 5.10 Sample Question 10 Results (Multiple-Choice)

Overall percentage correct and percentages correct within each achievement level range: 2001

Grade 12	Percentage correct within achievement-level intervals			
Overall percentage correct	Below *Basic* 293 and below*	*Basic* 294–324*	*Proficient* 325–354*	*Advanced* 355 and above*
68	54	82	95	***

*NAEP U.S. history composite scale range.
***Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

BEST COPY AVAILABLE

Grade 12 Sample Question 11:

Article I, Section 2, United States Constitution

"[The population of the states] shall be determined by adding to the whole number of free persons . . . three-fifths of all other persons."

An important debate led to the writing of this section of the Constitution. Identify the issue being debated.

Describe the northern position in this debate and explain why many northerners took it.

Describe the southern position in this debate and explain why many southerners took it.

Historical Theme:	Historical Period:
Change and Continuity in American Democracy: Ideas, Institutions, Practices, and Controversies	The Revolution and the New Nation (1763 to 1815)

Responses to this question were scored according to a four-level rubric as Inappropriate, Partial, Essential, or Complete

Table 5.11 Sample Question 11 Results (Extended Constructed-Response)

Overall percentage "Essential" or better and percentages "Essential" or better within each achievement level range: 2001

Grade 12	Percentage "Essential" or better within achievement-level intervals			
Overall percentage "Essential" or better	Below *Basic* 293 and below*	*Basic* 294–324*	*Proficient* 325–354*	*Advanced* 355 and above*
21	4	30	74	***

*NAEP U.S. history composite scale range.
***Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Responses scored "Complete" identified and explained the debate between the North and the South over the counting of slaves for purposes of representation, making explicit the relationship between counting the slaves and representation in Congress.

Sample "Complete" Response:

An important debate led to the writing of this section of the Constitution. Identify the issue being debated.

Wather or not blacks would be counted in the population.

Describe the northern position in this debate and explain why many northerners took it.

Northerners didn't want them to count cause the South would get more seats in the House of Reps

Describe the southern position in this debate and explain why many southerners took it.

They wanted them to count to get more votes in the House of Reps.

BEST COPY AVAILABLE

Responses scored "Essential" identified the debate without explicitly noting the desire of the southern states to maximize their voting power in the House of Representatives.

Sample "Essential" Response:

An important debate led to the writing of this section of the Constitution. Identify the issue being debated.

The issue is whether to count slaves in the population

Describe the northern position in this debate and explain why many northerners took it.

The northern position was not to count slaves because it would give the south more people

Describe the southern position in this debate and explain why many southerners took it.

The southern position was to count the slaves so they would have more people

BEST COPY AVAILABLE

Responses scored "Partial" identified the issue but did not make the northern and southern positions clear.

Sample "Partial" Response:

An important debate led to the writing of this section of the Constitution. Identify the issue being debated.

The issue being debated was the census of states population including slaves

Describe the northern position in this debate and explain why many northerners took it.

The north took this position for tax purposes

Describe the southern position in this debate and explain why many southerners took it.

The south took it as well for tax purposes.

BEST COPY AVAILABLE

Grade 12 Sample Question 12:

"In spite of the obvious advantages held by the North, the South was able to fight for four years and to achieve some real military successes. So while the North held most of the cards, the South had one or two aces up its sleeves."

Identify two of the "aces" (significant advantages) that the South had in the Civil War. Explain how these advantages helped the South.

Historical Theme:	Historical Period:
Change and Continuity in American Democracy: Ideas, Institutions, Practices, and Controversies	Crisis of the Union: Civil War and Reconstruction (1850 to 1877)

Responses to this question were scored according to a four-level rubric as Inappropriate, Partial, Essential, or Complete

Table 5.12 Sample Question 12 Results (Extended Constructed-Response)

Overall percentage "Essential" or better and percentages "Essential" or better within each achievement level range: 2001

Grade 12	Percentage "Essential" or better within achievement-level intervals			
Overall percentage "Essential" or better	*Below Basic* 293 and below*	*Basic* 294–324*	*Proficient* 325–354*	*Advanced* 355 and above*
39	17	62	88	***

*NAEP U.S. history composite scale range.
***Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Responses scored "Complete" identified two advantages held by the South and explained how each advantage identified aided the southern war effort.

Sample "Complete" Response:

Identify two of the "aces" (significant advantages) that the South had in the Civil War. Explain how these advantages helped the South.

The south did have a couple aces up their sleeves For one they were fighting the war in their home. They were familiar to the territory as well as having a shorter distance to reach supplies The other advantage they had was better military leaders. Robert E. Lee was asked to be the general of the North but he declined and became the South's general

BEST COPY AVAILABLE

Responses scored "Essential" identified one advantage the South had and explained how this advantage aided the southern war effort (or the responses identified two advantages but did not fully explain how both advantages aided the southern war effort).

Sample "Essential" Response:

Identify two of the "aces" (significant advantages) that the South had in the Civil War. Explain how these advantages helped the South.

Most of the Civil War was fought on southern land. Therefore the first of the two aces was that the southern soldiers knew the terrain better than the northern soldiers. The second of the two aces was that the southerners were used to the weather and the northerners were not.

Responses scored "Partial" identified one advantage the South had, but did not sufficiently explain how this advantage aided the southern war effort.

Sample "Partial" Response:

Identify two of the "aces" (significant advantages) that the South had in the Civil War. Explain how these advantages helped the South.

One significant advantage that the South had in the Civil War was that they were fighting on their own territory. This gave them sort of a 'home court' advantage.

BEST COPY AVAILABLE

Maps of Selected Item Descriptions on the NAEP U.S. History Scale—Grades 4, 8, and 12

The U.S. history performance of fourth-, eighth-, and twelfth-graders can be illustrated by maps that position item descriptions along the NAEP U.S. history scale where items are likely to be answered successfully by students.[1] The descriptions used on the maps focus on the U.S. history knowledge or skill needed to answer the question. For multiple-choice questions, the description indicates the knowledge or skill demonstrated by selection of the correct option; for constructed-response questions, the description takes into account the knowledge or skill specified by the different levels of scoring criteria for that question. Five of the questions described on the item maps are included among the sample questions in the preceding section. Each of these sample questions is identified as such on the item map.

Figures 5.1 through 5.3 are item maps for grades 4, 8, and 12, respectively. The map location for each question identifies where that question was answered successfully by at least 65 percent of the students for constructed-response questions and 74 percent of the students for four-option multiple-choice questions. For each question indicated on the map, students whose average score fell at or above the scale point had a higher probability of successfully answering the question, and students whose average score was below that scale point had a lower probability of successfully answering the question.

As an example of how to interpret the item maps, consider the multiple-choice question in figure 5.1 that maps at score point 245. As the description indicates, fourth-graders were required to "identify how railroad affected Chicago." As this was a four-option multiple-choice question, students with an average score at or above 245 had at least a 74 percent probability of answering the question correctly. Students with an average score below 245 had less than a 74 percent probability of doing so. This does not mean that all students with an average score of 245 or above always answered the question correctly, or that all students scoring below 245 always answered the question incorrectly. Rather, the item map indicates higher or lower probability of answering the question successfully depending on students' overall U.S. history ability as measured by the NAEP scale.

The three U.S. history achievement levels for a specific grade are indicated on the item map for that grade. It is important to note that, although the same 0–500 U.S. history scale is used at each grade, the achievement levels are grade specific and each achievement level begins at a different score point at each grade.

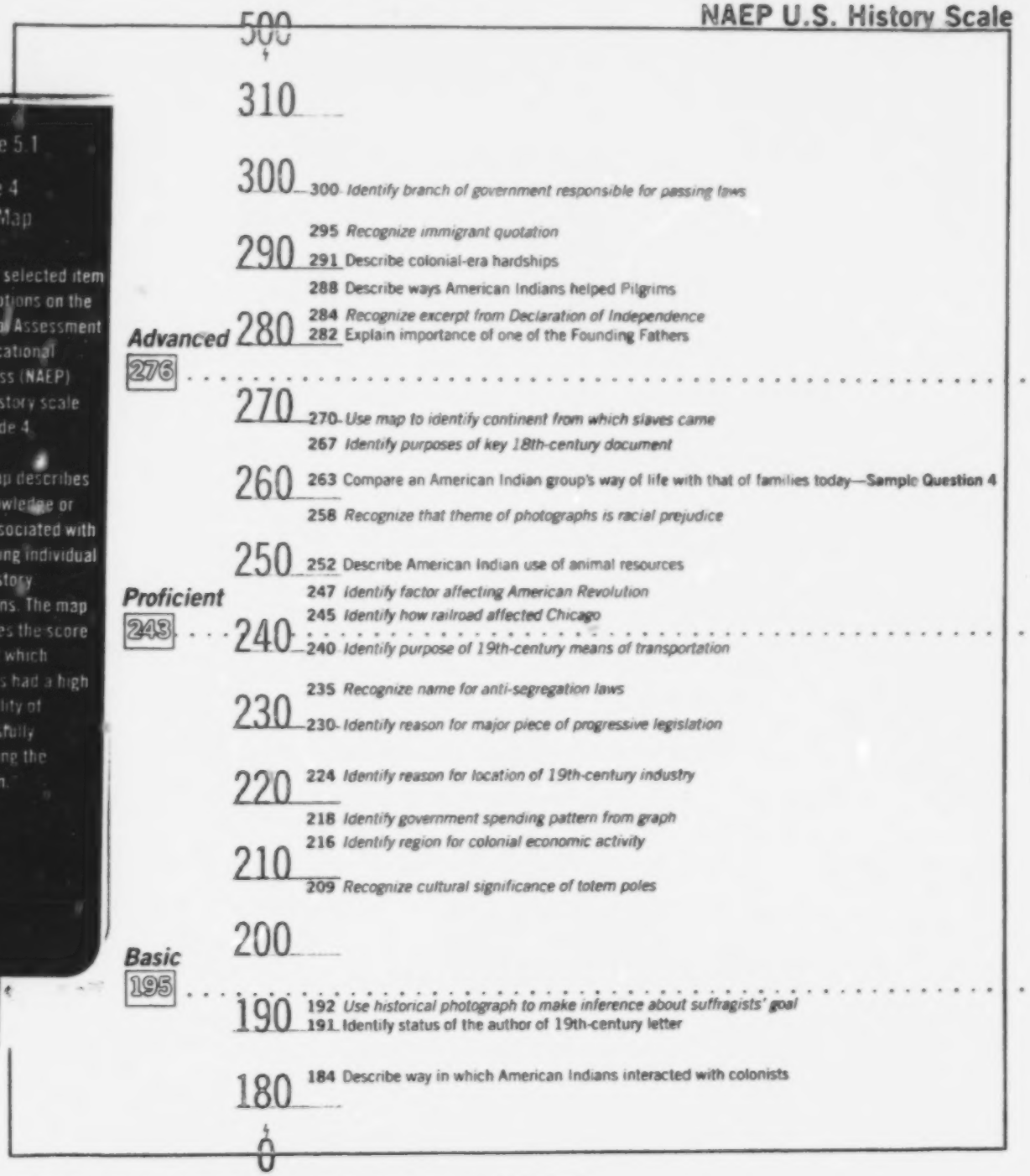
[1] Details on the procedures used to develop item maps are provided in appendix A.

Figure 5.1

Grade 4
Item Map

Map of selected item descriptions on the National Assessment of Educational Progress (NAEP) U. S. history scale for grade 4.

This map describes the knowledge or skill associated with answering individual U. S. history questions. The map identifies the score point at which students had a high probability of successfully answering the question.*


NAEP U.S. History Scale
500
310
300
300 Identify branch of government responsible for passing laws
295 Recognize immigrant quotation
290
291 Describe colonial-era hardships
288 Describe ways American Indians helped Pilgrims
284 Recognize excerpt from Declaration of Independence
280
282 Explain importance of one of the Founding Fathers
Advanced
276
270
270 Use map to identify continent from which slaves came
267 Identify purposes of key 18th-century document
260
263 Compare an American Indian group's way of life with that of families today—Sample Question 4
258 Recognize that theme of photographs is racial prejudice
250
252 Describe American Indian use of animal resources
247 Identify factor affecting American Revolution
Proficient
243
245 Identify how railroad affected Chicago
240
240 Identify purpose of 19th-century means of transportation
235 Recognize name for anti-segregation laws
230
230 Identify reason for major piece of progressive legislation
220
224 Identify reason for location of 19th-century industry
218 Identify government spending pattern from graph
216 Identify region for colonial economic activity
210
209 Recognize cultural significance of totem poles
200
Basic
195
190
192 Use historical photograph to make inference about suffragists' goal
191 Identify status of the author of 19th-century letter
180
184 Describe way in which American Indians interacted with colonists
0

NOTE: Regular type denotes a constructed-response question. Italic type denotes a multiple-choice question.

* Each grade 4 U. S. history question in the 2001 assessment was mapped onto the NAEP 0–500 U. S. history scale. The position of the question on the scale represents the scale score attained by students who had a 65 percent probability of successfully answering a constructed-response question, or a 74 percent probability of correctly answering a four-option multiple-choice question. Only selected questions are presented. Scale score ranges for U. S. history achievement levels are referenced on the map. For constructed-response questions, the question description represents students' performance at the scoring criteria level being mapped.

SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U. S. History Assessment.

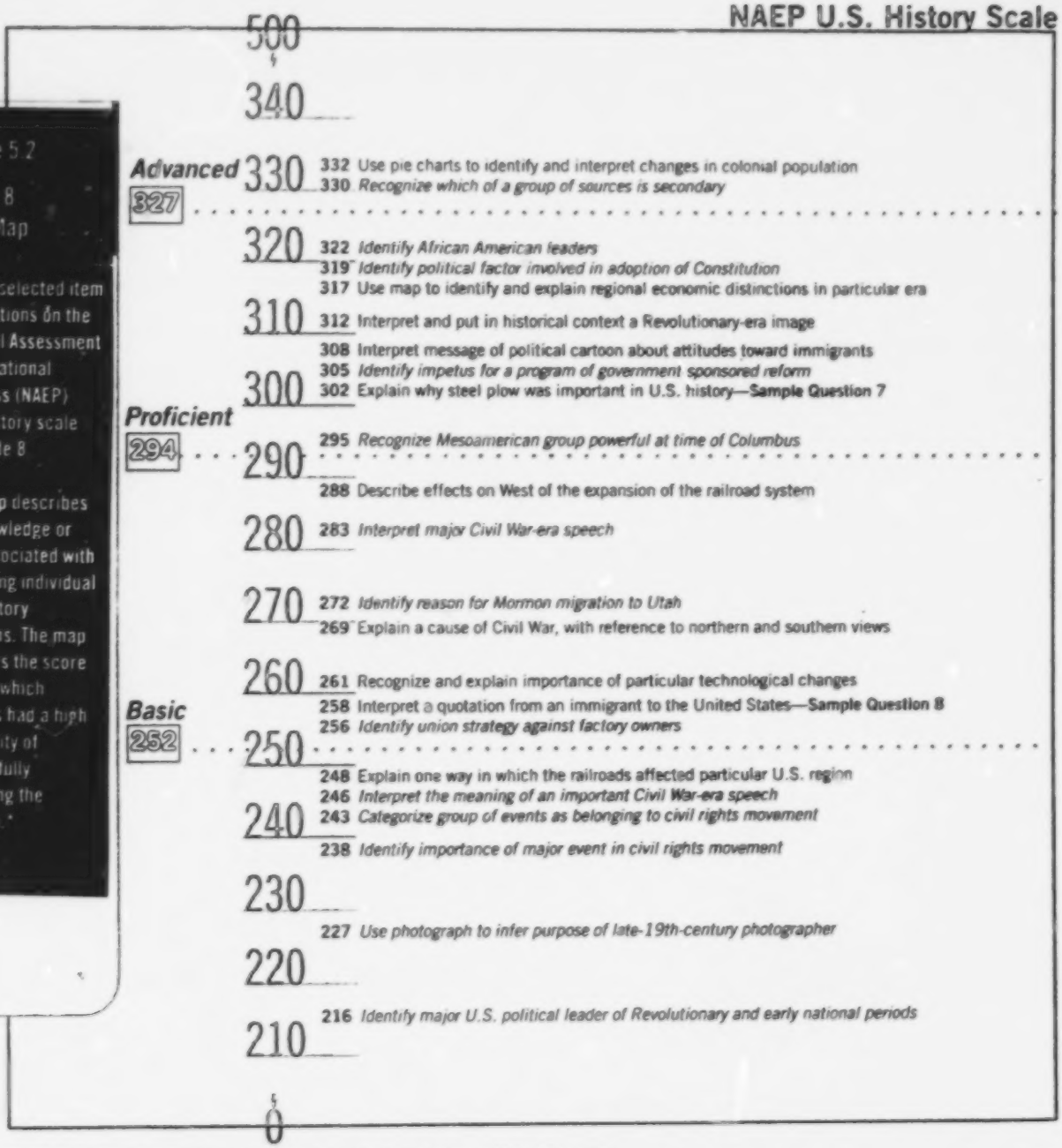
BEST COPY AVAILABLE

Figure 5.2

Grade 8
Item Map

Map of selected item descriptions on the National Assessment of Educational Progress (NAEP) U. S. history scale for grade 8

This map describes the knowledge or skill associated with answering individual U. S. history questions. The map identifies the score point at which students had a high probability of successfully answering the question.*


NAEP U.S. History Scale
500
340
Advanced
330
332 Use pie charts to identify and interpret changes in colonial population
330 Recognize which of a group of sources is secondary
327
320
322 Identify African American leaders
319 Identify political factor involved in adoption of Constitution
317 Use map to identify and explain regional economic distinctions in particular era
310
312 Interpret and put in historical context a Revolutionary-era image
308 Interpret message of political cartoon about attitudes toward immigrants
305 Identify impetus for a program of government sponsored reform
300
302 Explain why steel plow was important in U.S. history—Sample Question 7
Proficient
294
295 Recognize Mesoamerican group powerful at time of Columbus
290
288 Describe effects on West of the expansion of the railroad system
280
283 Interpret major Civil War-era speech
270
272 Identify reason for Mormon migration to Utah
269 Explain a cause of Civil War, with reference to northern and southern views
260
261 Recognize and explain importance of particular technological changes
258 Interpret a quotation from an immigrant to the United States—Sample Question 8
Basic
256 Identify union strategy against factory owners
252
250
248 Explain one way in which the railroads affected particular U.S. region
246 Interpret the meaning of an important Civil War-era speech
240
243 Categorize group of events as belonging to civil rights movement
238 Identify importance of major event in civil rights movement
230
227 Use photograph to infer purpose of late-19th-century photographer
220
216 Identify major U.S. political leader of Revolutionary and early national periods
210
0

NOTE: Regular type denotes a constructed-response question. Italic type denotes a multiple-choice question.

* Each grade 8 U. S. history question in the 2001 assessment was mapped onto the NAEP 0–500 U. S. history scale. The position of the question on the scale represents the scale score attained by students who had a 65 percent probability of successfully answering a constructed-response question, or a 74 percent probability of correctly answering a four-option multiple-choice question. Only selected questions are presented. Scale score ranges for U. S. history achievement levels are referenced on the map. For constructed-response questions, the question description represents students' performance at the scoring criteria level being mapped.

SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U. S. History Assessment.

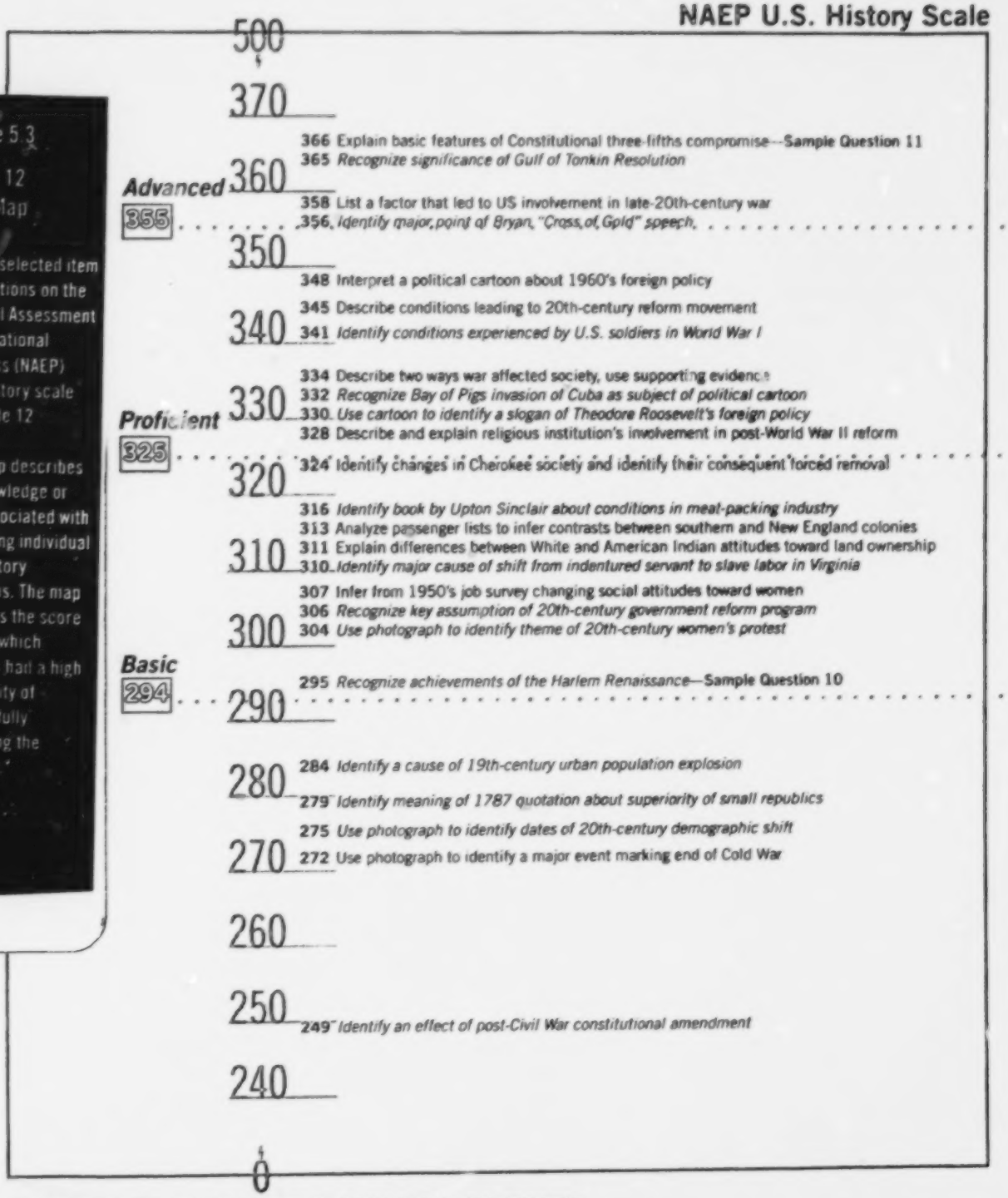
BEST COPY AVAILABLE

Figure 5.3

Grade 12
Item Map

Map of selected item descriptions on the National Assessment of Educational Progress (NAEP) U. S. history scale for grade 12

This map describes the knowledge or skill associated with answering individual U. S. history questions. The map identifies the score point at which students had a high probability of successfully answering the question.*


NAEP U.S. History Scale
500
370
366 Explain basic features of Constitutional three-fifths compromise---Sample Question 11
365 Recognize significance of Gulf of Tonkin Resolution
Advanced
360
358 List a factor that led to US involvement in late-20th-century war
355
356 Identify major point of Bryan, "Cross of Gold" speech
350
348 Interpret a political cartoon about 1960's foreign policy
345 Describe conditions leading to 20th-century reform movement
340
341 Identify conditions experienced by U.S. soldiers in World War I
334 Describe two ways war affected society, use supporting evidence
332 Recognize Bay of Pigs invasion of Cuba as subject of political cartoon
330
330 Use cartoon to identify a slogan of Theodore Roosevelt's foreign policy
Proficient
328 Describe and explain religious institution's involvement in post-World War II reform
325
324 Identify changes in Cherokee society and identify their consequent forced removal
320
316 Identify book by Upton Sinclair about conditions in meat-packing industry
313 Analyze passenger lists to infer contrasts between southern and New England colonies
311 Explain differences between White and American Indian attitudes toward land ownership
310
310 Identify major cause of shift from indentured servant to slave labor in Virginia
307 Infer from 1950's job survey changing social attitudes toward women
306 Recognize key assumption of 20th-century government reform program
300
304 Use photograph to identify theme of 20th-century women's protest
Basic
295 Recognize achievements of the Harlem Renaissance—Sample Question 10
294
290
284 Identify a cause of 19th-century urban population explosion
280
279 Identify meaning of 1787 quotation about superiority of small republics
275 Use photograph to identify dates of 20th-century demographic shift
270
272 Use photograph to identify a major event marking end of Cold War
260
250
249 Identify an effect of post-Civil War constitutional amendment
240
0

NOTE: Regular type denotes a constructed-response question. Italic type denotes a multiple-choice question.

* Each grade 12 U. S. history question in the 2001 assessment was mapped onto the NAEP 0–500 U. S. history scale. The position of the question on the scale represents the scale score attained by students who had a 65 percent probability of successfully answering a constructed-response question, or a 74 percent probability of correctly answering a four-option multiple-choice question. Only selected questions are presented. Scale score ranges for U. S. history achievement levels are referenced on the map. For constructed-response questions, the question description represents students' performance at the scoring criteria level being mapped.

SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U. S. History Assessment.

6

Classroom Contexts for Learning

This chapter presents information about practices that take place in and around the school that may affect the teaching and learning of U.S. history—the amount of time students spend in social studies classes, adherence to state and local standards for social studies education, classroom practices, and the use of technology. The information in this chapter is based on teachers' and students' responses to background questions administered as part of the NAEP 2001 U.S. history assessment. The percentage of students and average scale scores are presented for each contextual variable reported in order to examine the relationship between students' home and school experiences and their performance on the assessment. In interpreting these data, readers are reminded that the relationship between contextual variables and student performance is not necessarily causal. There are many factors that may play a role in a student's performance on NAEP.

Chapter Focus

How much time do fourth-grade teachers devote to social studies instruction?

How do classroom activities and computer use relate to student achievement?

Chapter Contents

Time Spent on Social Studies

State and Local Standards

Instructional Activities

Use of Technology

Time Spent on Social Studies

The *NAEP U.S. History Framework* that served as the blueprint for the 1994 and 2001 assessments recognized that most fourth-graders do not have a formal class in U.S. history. (Attention was paid in the framework to ensuring coverage of material that is likely to be addressed in fourth-grade classrooms, such as state history.) The majority of fourth-grade teachers, however, did report instructing their students regularly in social studies. More than one-half of fourth-graders had teachers who reported spending

BEST COPY AVAILABLE

between 61 and 180 minutes per week on social studies instruction in 2001. About one-fifth of the students had teachers who reported spending more than 180 minutes per week on social studies, and 14 percent had teachers who reported spending 60 minutes or less.

The results presented in table 6.1 show a generally positive relationship between the amount of time teachers reported spending on social studies instruction and fourth-graders' performance. Students whose teachers reported spending 60 minutes or less per week had lower scores than those whose teachers reported spending more than 60 minutes per week. Fourth-graders whose teachers spent more than 180 minutes per week on social studies had the highest average scores.

Table 6.1

Percentage of students and average U.S. history scale scores by teachers' reports on the amount of time spent on social studies in a typical week at grade 4: 2001

Grade 4

	2001
Less than 30 minutes	2 191
30 to 60 minutes	12 195
61 to 120 minutes	37 210
121 to 180 minutes	31 211
More than 180 minutes	19 218

Time Spent on Fourth Grade Social Studies

Students whose teachers reported spending more than 180 minutes a week on social studies scored highest

The percentage of students is listed first with the corresponding average scale score presented below.
NOTE: Percentages may not add to 100 due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

BEST COPY AVAILABLE

State and Local Standards

With the rise of the movement for standards-based reform in K-12 education, an increasing number of states have adopted standards for history or social studies.[1]

In 2001, teachers of fourth- and eighth-grade students were asked about the extent to which they used state or local standards in planning their history or social studies instruction. Table 6.2 presents the percentage of students and their average scores based on teachers' responses to this question. Only a small percentage of students had teachers who reported that there were no state or local standards that applied to teaching social studies (3 percent at grade 4, and 1 percent at grade 8). About two-thirds of the students had teachers who reported that standards were used to a large extent in planning instruction (63 percent at grade 4, and 69 percent at grade 8). There were, however, no statistically significant differences in students' performance at either grade 4 or grade 8, based on whether or not there were standards or on the extent to which teachers reported using standards for planning social studies instruction. Because state and local standards are diverse and are used in various ways, readers should interpret this data with caution.

1 Council of Chief State School Officers (2000). *Key state education policies on K-12 education* (table 13, p. 23). Washington DC: Author

Table 6.2

Percentage of students and average U.S. history scale scores by teachers' reports on the use of state/local standards in planning instruction at grades 4 and 8: 2001

Grades 4 & 8

	2001
Grade 4	
Not at all	2 212
Small extent	9 210
Moderate extent	23 206
Large extent	63 210
No standards for teaching social studies	3 224
Grade 8	
Not at all	2 274
Small extent	7 264
Moderate extent	21 266
Large extent	69 262
No standards for teaching social studies	1 276

Use of State/Local Standards in Planning Instruction

More than half of fourth- and eighth-grade students had teachers who used state/local standards to a large extent in planning instruction

The percentage of students is listed first with the corresponding average scale score presented below.
NOTE: Percentages may not add to 100 due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Instructional Activities

Fourth- and eighth-grade teachers whose students participated in the NAEP U.S. history assessments in 1994 and 2001 were asked a series of questions about the frequency with which they engaged their classes in certain types of instructional activities when teaching social studies or U.S. history. The results, presented in table 6.3, vary somewhat by grade level and show that most students had teachers who reported having them read from a textbook at least once or twice a week (88 percent at grade 4, and 90 percent at grade 8). In 2001, fourth-graders whose teachers asked them to read from a textbook on a daily basis had higher average scores than students whose teachers did so on a weekly or monthly basis. Although only a small percentage of fourth-graders had teachers

who reported never or hardly ever asking them to read from a textbook, the average scores for these students were not significantly different from those whose teachers did so more frequently. In contrast, eighth-graders whose teachers reported never or hardly ever having them read from a textbook had higher scores than their peers whose teachers engaged in this activity daily, weekly, or monthly. Here again, this involved a small percentage of students.

A large majority of fourth- and eighth-graders were in classes where teachers reported asking students to read extra materials that were not in the regular textbook, such as biographies or historical fiction, on at least an occasional basis (weekly or monthly). There were, however, no differences in students' performance at either grade, related to the frequency with which teachers reported having them read these extra materials.

Teachers were also asked about the frequency of their use of primary historical documents, letters, diaries, or essays written by historical figures. While the use of such texts as part of history or social studies instruction was not related to student performance at grade 4, there was a positive relationship associated with weekly use of these materials at grade 8. Eighth-graders whose teachers reported using primary sources on a weekly basis had higher average scores than those whose teachers reported doing so monthly or never.

A question that asked teachers about the frequency with which they engaged their students in writing reports revealed different performance patterns at grades 4 and 8. The average scores of fourth-graders whose teachers asked them to write reports weekly were lower than the scores of students whose teachers did so less frequently. There was no relationship between the performance of eighth-graders and the frequency of writing reports.

On the whole there has been little change between 1994 and 2001 in the percentage of students whose teachers reported various amounts of time spent in particular instructional activities. There were some exceptions to this pattern of stability, however. At grade 4, the percentage of students with teachers who reported reading extra material once or twice a week increased from 33 percent in 1994 to 44 percent in 2001, while the percentage with teachers reporting doing so once or twice a month decreased. The percentage of fourth-graders with teachers who used historical documents once or twice a month also increased, while the percentage with teachers who never or hardly ever used them decreased. At grade 8, the percentage of students whose teachers reported having them read extra material every day increased from 3 percent in 1994 to 7 percent in 2001. The percentage of eighth-graders whose teachers never used primary historical documents decreased from 1994 to 2001.

Table 6.3

Percentage of students and average U.S. history scale scores by teachers' reports on frequency of classroom activities at grades 4 and 8: 1994 and 2001

Grade 4

Frequency of Fourth- and Eighth-Grade Classroom Activities

Daily reading from a textbook was associated with higher scores than reading on a weekly or monthly basis.

	1994	2001
Reading material from a textbook		
Almost every day	43 207	41 214
Once or twice a week	44 204	47 207
Once or twice a month	8 204	7 202
Never or hardly ever	5 204	5 209
Reading extra material not in the regular textbook		
Almost every day	6 208	9 210
Once or twice a week	33 205	44 * 211
Once or twice a month	46 204	35 * 208
Never or hardly ever	15 208	12 208
Using primary historical documents		
Almost every day	1 ***	1 ***
Once or twice a week	8 201	11 207
Once or twice a month	29 208	39 * 212
Never or hardly ever	62 205	48 * 208
Writing a report		
Almost every day	*** ***	1 194
Once or twice a week	6 188	6 198
Once or twice a month	63 207	59 210
Never or hardly ever	31 205	34 210

See footnotes at end of table. ▶

Table 6.3 (continued)

Percentage of students and average U.S. history scale scores by teachers' reports on frequency of classroom activities at grades 4 and 8: 1994 and 2001

Grade 8

Frequency of Fourth- and Eighth-Grade Classroom Activities

	1994	2001
Reading material from a textbook		
Almost every day	45 259	45 264
Once or twice a week	42 259	45 262
Once or twice a month	8 266	7 262
Never or hardly ever	5 265	3 275
Reading extra material not in the regular textbook		
Almost every day	3 254	7 * 265
Once or twice a week	32 258	37 261
Once or twice a month	47 263	44 264
Never or hardly ever	17 258	13 264
Using primary historical documents		
Almost every day	2 268	4 264
Once or twice a week	20 260	27 267
Once or twice a month	55 261	54 262
Never or hardly ever	23 258	16 * 259
Writing a report		
Almost every day	# ***	1 255
Once or twice a week	4 256	7 266
Once or twice a month	66 261	66 263
Never or hardly ever	30 259	27 263

Weekly use of primary documents was associated with higher scores than less frequent use

The percentage of students is listed first with the corresponding average scale score presented below.
*Significantly different from 1994.
*** Sample size is insufficient to permit a reliable estimate.
# Percentage is between 0.0 and 0.5.
NOTE: Percentages may not add to 100 due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

Table 6.4 presents the data for questions asked of twelfth-grade students that were similar to those asked of fourth- and eighth-grade teachers. A large majority of students reported reading from a textbook on a daily or weekly basis and had higher average scores than those who did so only a few times a year or never. Students who reported that they never read extra material, such as biographies or historical stories, not in the regular textbook had lower average scores than students who did so a few times a year or more. However, there was no statistically significant difference in the performance of students who read such extra material as frequently as every day compared to others who used it to a more limited extent, including those who did so as infrequently as a few times a year. Both twelfth-graders who reported never using letters, diaries, or essays written by historical people and those who reported doing so on a daily basis had lower average scores than students who reported engaging in this activity on a more moderate basis—weekly, monthly, or yearly. A similar pattern was evident in the results presented for writing reports.

A general pattern noticeable between 1994 and 2001 was that an increasing percentage of twelfth-graders reported regularly engaging in the instructional activities that they were asked about. The percentage of students who reported daily reading from a textbook increased from 40 percent in 1994 to 44 percent in 2001. There were also higher percentages of twelfth-graders who reported using historical documentation and writing reports daily, weekly, and monthly in 2001 than in 1994.

Table 6.4

Percentage of students and average U.S. history scale scores by students' reports on frequency of classroom activities at grade 12: 1994 and 2001

Grade 12

Frequency of Twelfth-Grade Classroom Activities

	1994	2001
Read material from a textbook		
About every day	40	44 *
	289	290
Once or twice a week	40	38
	289	289
Once or twice a month	9	8
	284	283
A few times a year	6	6
	278	276
Never	6	5 *
	268	270
Read extra material not in the regular textbook		
About every day	9	10
	288	290
Once or twice a week	30	31
	289	291
Once or twice a month	24	25
	291	290
A few times a year	18	17
	288	289
Never	18	17
	274	276

Students who reported never reading extra material scored lowest.

See footnotes at end of table. ▷

Table 6.4 (continued)

Percentage of students and average U.S. history scale scores by students' reports on frequency of classroom activities at grade 12: 1994 and 2001

Grade 12

	1994	2001
Use letters, diaries, or essays written by historical people		
About every day	4 280	5 * 280
Once or twice a week	14 292	18 * 290
Once or twice a month	24 291	26 * 291
A few times a year	26 291	26 292
Never	32 279	25 * 279
Write a report		
About every day	2 267	3 * 271
Once or twice a week	9 279	14 * 288
Once or twice a month	35 287	41 * 290
A few times a year	40 293	34 * 290
Never	14 278	8 * 271

The percentage of students is listed first with the corresponding average scale score presented below.
*Significantly different from 1994.
NOTE: Percentages may not add to 100 due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

Use of Technology

Computer use in history and social studies education is an area of rapidly growing interest on the part of history educators.[2] Fourth-, eighth-, and twelfth-grade students who participated in the 2001 assessment were asked a series of questions related to their use of computers for history and social studies both at school and at home. The results presented in tables 6.5 and 6.6 show the relationship between students' performance and their responses to three questions about their use of computers.

Relatively few students reported using a computer for social studies or history. Only about one-quarter of the fourth-grade students reported using computers at school for social studies at least once every few weeks. This proportion increased to about one-third of the students at grade 8, and remained at about one-third among students at grade 12.

The data presented in table 6.5 for grade 4 and table 6.6 for grades 8 and 12, show that the reported frequency related to general usage of computers at school for history or social studies appeared to have a negative association with students' performance at all three grades. Students in grades 4, 8, and 12 who reported daily use of computers at school for social studies all had lower average scores than those who reported less frequent use. There was not, however, a consistently negative association between computer use and students' performance. A positive association with performance was evident for using computers specifically for conducting research and for writing reports at grades 8 and 12. At grades 8 and 12, students who used a CD or the Internet for research projects or used the computer to write reports more frequently had higher average scores than their peers who did so less frequently. This pattern did not hold true at grade 4, however, where there was no significant difference between the performance of those students who reported using a CD or the Internet for research projects or to write reports and those who reported that they did not.

2 Martorella, P. H. (Ed.). (1997). *Interactive technologies and the social studies: Emerging issues and applications.* Albany, NY: State University of New York Press.

Table 6.5

Percentage of students and average U.S. history scale scores by students' reports on computer use at grade 4: 2001

Grade 4

Fourth Grade Computer Use

More frequent general computer use at school was associated with lower scores.

	2001
Use computers at school for social studies	
Every day	3 (167)
Two or three times a week	5 186
Once a week	7 197
Once every few weeks	10 212
Never or hardly ever	74 214
Do research projects using a CD or the Internet	
Yes	46 211
No	54 208
Use computer to write reports	
Yes	52 209
No	48 210

The percentage of students is listed first with the corresponding average scale score presented below.
NOTE: Percentages may not add to 100 due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Table 6.6

Percentage of students and average U.S. history scale scores by students' reports on computer use at grades 8 and 12: 2001

Grade 8

Eighth- and Twelfth-Grade Computer Use

	2001
Use computers at school for social studies	
Every day	1 239
Two or three times a week	5 252
Once a week	9 261
Once every few weeks	21 268
Never or hardly ever	64 263
Do research projects using a CD or the Internet	
Not at all	26 253
Small extent	31 262
Moderate extent	26 267
Large extent	16 272
Write reports on the computer	
Not at all	18 253
Small extent	30 260
Moderate extent	29 266
Large extent	23 270

More extensive computer use for research projects was associated with higher scores.

See footnotes at end of table. >

Table 6.6 (continued)

Percentage of students and average U.S. history scale scores by students' reports on computer use at grades 8 and 12: 2001

Grade 12

	2001
Use computers at school for studying history	
Every day	2 265
Two or three times a week	6 277
Once a week	7 280
Once every few weeks	16 291
Never or hardly ever	42 289
Haven't studied history this year	27 289
Do research projects using a CD or the Internet	
Not at all	23 274
Small extent	33 286
Moderate extent	29 294
Large extent	15 300
Write reports on the computer	
Not at all	14 271
Small extent	27 281
Moderate extent	33 290
Large extent	26 300

The percentage of students is listed first with the corresponding average scale score presented below.
NOTE: Percentages may not add to 100 due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Appendix A
Overview of Procedures Used for the NAEP 2001 U. S. History Assessment

This appendix provides an overview of the NAEP 2001 history assessment's primary components—framework, development, administration, scoring, and analysis. A more extensive review of the procedures and methods used in the history assessment will be included in the forthcoming *NAEP 2001 Technical Report.*

Appendix Contents

The Assessment
The Samples
Students with Disabilities (SD) and Limited English Proficient (LEP) Students
Data Collection
Data Analysis
NAEP Reporting Groups
Cautions in Interpretations

Appendix Focus

Technical aspects of the NAEP 2001 history assessment

The NAEP 2001 History Assessment

The National Assessment Governing Board (NAGB), created by Congress in 1988, is responsible for formulating policy for NAEP. The NAGB is specifically charged with developing assessment objectives and test specifications through a national consensus approach. That consensus approach results in the development of an assessment framework. The design of the NAEP 2001 U.S. history assessment followed the guidelines provided in the framework developed for the 1994 assessment.[1]

The framework underlying both the NAEP 1994 and 2001 assessments reflects current consensus among educators and researchers about the study of U.S. history. Developing this framework and the specifications that guided development of the assessment involved the critical input of hundreds of individuals across the country, including representatives of national education

[1] National Assessment Governing Board (1994). *U.S. history framework for the 1994 National Assessment of Educational Progress.* Washington, DC: Author.

BEST COPY AVAILABLE

organizations, teachers, parents, policymakers, business leaders, and the interested general public. This consensus process was managed by the Council of Chief State School Officers for NAGB.

The assessment framework specified not only the particular aspects of U.S. history to be measured (see chapter 1 for a description of these aspects), but also the percentage of assessment questions that should be devoted to each. The target percentage distributions of historical themes, as specified in the framework, along with the actual percentage distributions in the 1994 and 2001 assessments, are presented in table A.1. Notice that these percentages shift from grade 4 to grade 12 to reflect the shift in curricular emphasis as students move from the fourth to the twelfth grade. For example, the emphasis on "the changing role of America in the world" grows at each successive grade level.

It should also be noted that the actual content of the assessment has varied somewhat in both 1994 and 2001 from the targeted distribution. At grades 8 and 12, these variances are in part explained by the use of "theme blocks," which focus on particular historical themes and allow students to respond to a range of primary sources more extensively than would be possible in blocks that include the full spectrum of themes. In addition, it is evident from table A.1 that some variance exists between the actual distribution of questions among the themes in 1994 and the actual distribution in 2001. These variances exist because seven new blocks were introduced in 2001 to replace blocks that were released to the public. Within the new blocks, the distribution of items differed somewhat from the distribution within the blocks that they replaced.

Table A.1 Distribution of Questions

Target and actual percentage distribution of questions by historical theme, grades 4, 8, and 12: 1994 and 2001

	Grade 4			Grade 8			Grade 12		
Historical Themes	Target	Actual 1994	Actual 2001	Target	Actual 1994	Actual 2001	Target	Actual 1994	Actual 2001
Change and Continuity in American Democracy: Ideas, Institutions, Practices, and Controversies	25	24	25	30	28	30	25	29	28
The Gathering and Interactions of Peoples, Cultures, and Ideas	35	32	32	30	30	32	25	23	26
Economic and Technological Changes and Their Relation to Society, Ideas, and the Environment	25	25	32	20	23	25	25	26	22
The Changing Role of America in the World	15	19	12	20	19	13	25	21	25

SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

The Assessment Design

Each student who participated in the U.S. history assessment received a booklet containing three or four sections: a set of general background questions, a set of subject-specific background questions dealing largely with the student's use of technology, and one or two sets, or "blocks," of cognitive questions assessing knowledge and skills in U.S. history as outlined in the framework. At grade 4, only 25-minute blocks were used. At grades 8 and 12, students were given either two 25-minute blocks or one 50-minute block. The 50-minute blocks administered at grades 8 and 12 (one at each grade) included extended constructed-response questions requiring students to synthesize elements from various primary sources.

At grade 4 a total of six blocks of cognitive questions were given, while at grades 8 and 12 nine blocks were administered.[2] Some of the blocks at each grade level (three at grade 4, six at grade 8, and six at grade 12) were carried forward from the 1994 assessment to the 2001 assessment to allow for the measurement of change across time. Each block consisted of both multiple-choice and constructed-response questions. Short constructed-response questions required a few sentences for an answer, while extended constructed-response questions generally required a paragraph or more. It was expected that

[2] These blocks were distributed across the student booklets in a Balanced Incomplete Block (BIB) design that is described later in this section.

students could adequately answer the short constructed-response questions in about two to three minutes and the extended constructed-response questions in about five minutes.

The data in table A.2 display the number of questions by type and by grade level for the 1994 and 2001 assessments. Some of these questions were used at more than one grade level; thus, the sum of the questions that appear at each grade level is greater than the total number of unique questions. The total number of questions at each grade level was slightly smaller in 2001 than in 1994 because in 2001 there were slightly fewer extended constructed-response questions at grades 4 and 12, and fewer multiple-choice questions at grade 8. This decrease in the total number of questions simply reflects the fact that the new blocks that replaced blocks released to the public from the 1994 assessment contained slightly fewer questions. It should be noted that these variations across years do not affect the ability of NAEP to report changes in students' performance across years since the estimated changes are based on the presence of blocks that were common to both assessment years.

Table A.2 Distribution of Questions by Question Type

Distribution of questions administered by question type, grades 4, 8, and 12: 1994 and 2001

	Grade 4		Grade 8		Grade 12	
	1994	2001	1994	2001	1994	2001
Multiple-choice	63	63	102	99	103	106
Short constructed-response	26	28	37	39	33	35
Extended constructed-response	6	3	12	7	19	13
Total	95	94	151	145	155	154

SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

BEST COPY AVAILABLE

The assessment design allowed for maximum coverage of U.S. history content at grades 4, 8, and 12, while minimizing the time burden for any one student. This was accomplished through the use of matrix sampling of questions, in which representative samples of students took various portions of the entire pool of assessment questions. The aggregate results across the entire assessment allowed for broad reporting of the U.S. history abilities for the targeted population.

In addition to matrix sampling, the assessment design utilized a procedure for distributing booklets that controlled for position and context effects. Students received different blocks of questions in their booklets according to a procedure called "Balanced Incomplete Block (BIB) spiraling." This procedure assigns blocks of questions so that every block appears in the first or second position within a booklet an equal number of times. Every block of questions is paired with every other block, with the exception of the 50-minute theme block, which appears on its own without another block of cognitive questions. The spiraling aspect of this procedure cycles the booklets for administration, so that typically only a few students in any assessment session receive the same booklet. This design allows for some balancing of the impact of context and fatigue effects to be measured and reported, but makes allowance for the difficulties of administering the 50-minute blocks.[3]

In addition to the student assessment booklets, three other instruments provided data relating to the assessment: a teacher questionnaire, a school questionnaire, and a Students with Disabilities/Limited-English Proficiency (SD and/or LEP) questionnaire. The teacher questionnaire was administered to the history or social studies teachers of fourth- and eighth-grade students participating in the assessment. The questionnaire consisted of three sections and took approximately 20 minutes to complete. The first section focused on the teacher's general background and experience; the second section on computer resources available in the school; and the third section on classroom information about social studies instruction.

The school characteristics and policy questionnaire was given to the principal or other administrator in each participating school and took about 20 minutes to complete. The questions asked about school policies, programs, facilities, and the demographic composition and background of the student body.

The SD and/or LEP student questionnaire was completed by a school staff member knowledgeable about those students who were selected to participate in the assessment and who were identified as: 1) having an Individualized Education Program (IEP) or equivalent program (for reasons other than being gifted and talented) or 2) being limited English

[3] For further details on the booklet design, see the forthcoming *NAEP 2001 Technical Report*.

proficient (LEP). A questionnaire was completed for each SD and/or LEP student sampled regardless of whether the student participated in the assessment. Each questionnaire took approximately 3 minutes to complete and asked about the student and the special programs in which he or she participated.

National Sample

The national results presented in this report are based on a nationally representative probability sample of fourth-, eighth-, and twelfth-grade students. The sample was chosen using a multistage design that involved sampling students from selected schools within selected geographic areas across the country. The sample design had the following stages:

1) selection of geographic areas (a county, group of counties, or metropolitan statistical area);

2) selection of schools (public and nonpublic) within the selected areas; and

3) selection of students within selected schools.

Each selected school that participated in the assessment and each student assessed represents a portion of the population of interest. Sampling weights are needed to make valid inferences between the student samples and the respective populations from which they were drawn. Sampling weights account for disproportionate representation due to the oversampling of students who attend schools with high concentrations of Black and/or Hispanic students and students who attend nonpublic schools. Among other uses, sampling weights also account for lower sampling rates for very small schools and are used to adjust for school and student nonresponse.[4]

Unlike the 1994 national assessment, a special feature of the 2001 national assessment was the collection of data from samples of students where assessment accommodations for special-needs students were not permitted and from samples of students where accommodations for special-needs students were permitted. NAEP inclusion rules were applied, and accommodations were offered only when a student had an Individualized Education Program (IEP) because of a disability and/or was identified as being a limited English proficient student (LEP); all other students were asked to participate in the assessment under standard conditions.

Table A.3 shows the number of students included in the national samples for the NAEP 1994 and 2001 history assessments at each grade level. For the 2001 assessment, the table includes the number of students in the sample where accommodations were not permitted and the number of students in the sample where accommodations were permitted. The table shows that the same non-SD and/or LEP students were included in both samples in 2001; only the SD and/or LEP students differed between the two samples. The 1994 design differed somewhat in that the SD and/or LEP students were assessed in standard conditions and accommodations were not permitted.

[4] Additional details regarding the design and structure of the national and state samples will be included in the forthcoming *NAEP 2001 Technical Report*. In addition, the reader may consult the *NAEP 2000 Technical Report* for a discussion of sampling procedures that are mostly common to all NAEP assessments.

Table A.3 National Student Sample Size

National student sample size by type of results, grades 4, 8, and 12: 1994 and 2001

	1994	2001	
	Accommodations not permitted sample	Accommodations not permitted sample	Accommodations permitted sample
Grade 4			
Non SD/LEP students assessed	5,067	6,846	
SD/LEP students assessed without accommodations	432	581	504
SD/LEP students assessed with accommodations	NA	NA	359
Total students assessed	5,499	7,027	7,309
Grade 8			
Non SD/LEP students assessed	8,227	10,321	
SD/LEP students assessed without accommodations	540	918	863
SD/LEP students assessed with accommodations	NA	NA	569
Total students assessed	8,767	11,239	11,753
Grade 12			
Non SD/LEP students assessed	7,427	10,658	
SD/LEP students assessed without accommodations	391	658	566
SD/LEP students assessed with accommodations	NA	NA	253
Total students assessed	7,818	11,316	11,477

SD = Students with Disabilities.
LEP = Limited English Proficient students.
NA = Not applicable. No accommodations were permitted in this sample.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

BEST COPY AVAILABLE

Table A.4 provides a summary of the national school and student participation rates for the U.S. history assessment samples where accommodations were not permitted and where accommodations were permitted. Participation rates are presented for public and nonpublic schools, individually and combined. The first rate is the weighted percentage of schools participating in the assessment before substitution of demographically similar schools.[5] This rate is based only on the number of schools that were initially selected for the assessment. The numerator of this rate is the sum of the number of students represented by each initially selected school that participated in the assessment. The denominator is the sum of the number of students represented by each of the initially selected schools that had eligible students enrolled.

The second school participation rate is the weighted participation rate after substitution. The numerator of this rate is the sum of the number of students represented by each of the participating schools, whether originally selected or selected as a substitute for a school that chose not to participate. The denominator is the same as that for the weighted participation rate for the initial sample. Because of the common denominators, the weighted participation rate after substitution is at least as great as the weighted participation rate before substitution.

Also presented in table A.4 are weighted student participation rates. The numerator of this rate is the sum across all students assessed (in either an initial session or a makeup session) of the number of students that each represents. The denominator of this rate is the sum across all eligible sampled students in participating schools of the number of students that each represents. The overall participation rates take into account the weighted percentage of school participation before or after substitution and the weighted percentage of student participation after makeup sessions.

For the grade 12 national sample, where school and student response rates did not meet NCES standards, an extensive analysis was conducted that examined, among other factors, the potential for nonresponse bias at both the school and student level. No evidence of any significant potential for either school or student nonresponse bias was found. Results of these analyses, as well as nonresponse bias analyses for the grade 4 and grade 8 national samples, will be included in the forthcoming *NAEP 2001 Technical Report*.

BEST COPY AVAILABLE

[5] The initial base sampling weights were used in weighting the percentages of participating schools and students. An attempt was made to preselect (before field processes began) a maximum of two substitute schools for each sampled public school (one in-district and one out-of-district) and each sampled Catholic school, and one for each sampled nonpublic school (other than Catholic). To minimize bias, a substitute school resembled the original selection as much as possible on affiliation, estimated number of grade-eligible students, and minority composition.

Table A.4 Participation Rates

National school and student participation rates for public schools, nonpublic schools, and public and nonpublic schools combined, grades 4, 8, and 12: 2001

	Weighted school participation			Samples where accommodations were not permitted				Samples where accommodations were permitted			
				Student participation		Overall participation rate		Student participation		Overall participation rate	
	Percentage before substitution	Percentage after substitution	Total number of schools	Weighted percentage student participation	Total number of students assessed	Before substitution	After substitution	Weighted percentage student participation	Total number of students assessed	Before substitution	After substitution
Grade 4											
Public	83	88	276	96	5,978	80	84	96	6,266	80	84
Nonpublic	83	91	89	97	1,049	81	88	97	1,043	81	88
Combined	83	88	365	96	7,027	80	85	96	7,309	80	85
Grade 8											
Public	79	87	259	93	9,694	73	81	93	10,180	73	81
Nonpublic	84	88	110	96	1,561	81	84	96	1,582	80	84
Combined	79	87	369	93	11,255	74	81	93	11,762	74	81
Grade 12											
Public	73	80	311	77	10,051	56	62	76	10,220	56	61
Nonpublic	67	77	63	90	1,265	61	70	90	1,257	61	70
Combined	72	80	374	78	11,316	56	62	77	11,477	56	62

SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progres (NAEP), 2001 U. S. History Assessment.

Students with Disabilities (SD) and/or Limited English Proficient (LEP) Students

It is NAEP's intent to assess all selected students from the target population. Therefore, every effort is made to ensure that all selected students who are capable of participating in the assessment are assessed. Some students sampled for participation in NAEP can be excluded from the sample according to carefully defined criteria. These criteria were revised in 1996 to communicate more clearly a presumption of inclusion except under special circumstances. According to these criteria, students with Individualized Education Programs (IEPs) were to be included in the NAEP assessment except in the following cases:

1) The school's IEP team determined that the student could not participate, OR,

2) The student's cognitive functioning was so severely impaired that she or he could not participate, OR,

3) The student's IEP required that the student had to be tested with an accommodation or adaptation and that the student could not demonstrate his or her knowledge without that accommodation.[6]

[6] As described in the following section, a second sample in the 2001 national assessments was assessed that included students who required and were provided with accommodations.

All LEP students receiving academic instruction in English for three years or more were to be included in the assessment. Those LEP students receiving instruction in English for fewer than three years were to be included unless school staff judged them to be incapable of participating in the assessment in English.

Participation of SD and/or LEP Students in the NAEP Samples

Testing all sampled students is the best way for NAEP to ensure that the statistics generated by the assessment are as representative as possible of the performance of the entire national population and the populations of participating jurisdictions. However, all groups of students include certain proportions that cannot be tested in large-scale assessments (such as students who have profound mental disabilities), or who can only be tested through the use of "accommodations" such as extra time, one-on-one administration, or use of magnifying equipment.

Some students with disabilities and some LEP students cannot show on a test what they know and can do unless they are provided accommodations. When such accommodations are not allowed, students requiring such adjustments are often excluded from large-scale assessments such as NAEP. This phenomenon has become more common in the last decade and gained momentum with the passage of the Individuals with Disabilities Education Act (IDEA), which led schools and states to identify increasing proportions of students as needing accommodations on assessments to best show what they know and can do.[7] Furthermore, Section 504 of the Rehabilitation Act of 1973 requires that, when students with disabilities are tested, schools must provide them with appropriate accommodations so that the test results accurately reflect what the students know and are able to do.[8] In addition, as the proportion of limited English proficient students in the population has increased, some states have started offering accommodations, such as translated versions of assessments or the use of bilingual dictionaries as part of assessments.

Before 1996, NAEP did not allow any testing under nonstandard conditions (i.e., accommodations were not permitted). At that time, NAEP samples were able to include almost all sampled students in "standard" assessment sessions. However, as the influence of IDEA grew more widespread, the failure to provide accommodations led to increasing levels of exclusion in the assessment. Such increases posed two threats to the program: 1) they threatened the stability of trend lines (because excluding more students in one year than the next might lead to apparent rather than real gains), and 2) they made NAEP samples less than optimally representative of target populations.

NAEP reacted to this challenge by adopting a multipart strategy. It became clear that, to ensure that NAEP samples were as inclusive as possible, the program had to move toward allowing the same assessment accommodations that were afforded students in state and district

---

7 Office of Special Education Programs (1997). *Nineteenth annual report to Congress on the implementation of the individuals with disabilities education act.* Washington, DC: U. S. Department of Education.

8 Section 504 of the Rehabilitation Act of 1973 is a civil rights law designed to prohibit discrimination on the basis of disability in programs and activities, including education, that receive federal financial assistance.

testing programs. However, allowing accommodations represents a change in testing conditions that may affect measurement of changes over time. Therefore, beginning with the 1996 national assessments and the 1998 state assessments, NAEP has assessed a series of parallel samples of students. In one set of samples, testing accommodations were not permitted; this has allowed NAEP to maintain the measurement of achievement trends. In addition to the samples where accommodations were not permitted, parallel samples in which accommodations were permitted were also assessed. By having two overlapping samples and two sets of related data points, NAEP could meet two core program goals.[9] First, data trends could be maintained. Second, parallel trend lines could be set in ways that ensure that in future years the program will be able to use the most inclusive practices possible and mirror the procedures used by most state and district assessments. Beginning in 2002, NAEP will use only the more inclusive samples in which assessment accommodations are permitted.

In U.S. history, national data from 1994 and 2001 are reported for the sample in which accommodations were not permitted. National data for the second sample, in which accommodations were permitted, are reported at all grades for 2001 only.

In order to make it possible to evaluate the impact of increasing exclusion rates, data on exclusion in both assessment years are included in this appendix. Since the exclusion rates may affect average scale scores, readers should consider the magnitude of exclusion rate changes when interpreting score changes.

Percentages of students with disabilities (SD) and/or limited English proficient (LEP) students for the national sample where accommodations were not permitted are presented in table A.5. The data in this table include the percentages of students *identified* as SD and/or LEP, the percentage of students *excluded*, and the percentage of *assessed* SD and/or LEP students. Percentages of these students in the national sample where accommodations were permitted are presented in table A.6. The data in this table include the percentages of students *identified* as SD and/or LEP, the percentage of students *excluded*, the percentage of *assessed* SD and/or LEP students, the percentage *assessed without accommodations*, and the percentage *assessed with accommodations*.

In the 2001 accommodations-not-permitted national sample, 7 percent of students at grade 4, 8 percent of students at grade 8, and 4 percent of students at grade 12 were excluded from the assessment. The comparable percentages in the 2001 accommodations-permitted national sample were 3 percent at grades 4 and 8, and 2 percent at grade 12, respectively. This comparison would suggest that allowing accommodations did help to decrease the percentage of students excluded from the assessment.

[9] The two samples are described as "overlapping" because in 2001 the same group of non-SD and/or LEP students were included in both samples.

Table A.5 Students Identified as SD and/or LEP Where Accommodations Were Not Permitted

Percentage of students identified as SD and/or LEP where accommodations were not permitted, grades 4, 8, and 12: 1994 and 2001

	1994		2001	
	Number of students sampled	Weighted percentage of students	Number of students sampled	Weighted percentage of students
Grade 4				
SD and/or LEP students				
Identified	1,457	13	1,059	16
Excluded	1,025	5	478	7
Assessed	432	8	581	9
SD students only				
Identified	961	10	582	10
Excluded	685	4	346	5
Assessed	276	5	236	5
LEP students only				
Identified	531	4	521	6
Excluded	368	1	159	2
Assessed	163	2	362	4
Grade 8				
SD and/or LEP students				
Identified	1,818	11	1,727	16
Excluded	1,278	5	809	8
Assessed	540	6	918	8
SD students only				
Identified	1,358	8	1,197	12
Excluded	979	4	671	7
Assessed	379	5	526	5
LEP students only				
Identified	486	2	605	4
Excluded	323	1	187	1
Assessed	163	1	418	3
Grade 12				
SD and/or LEP students				
Identified	1,339	8	1,336	11
Excluded	948	3	678	4
Assessed	391	5	658	6
SD students only				
Identified	1,013	6	913	8
Excluded	776	3	567	4
Assessed	237	3	346	4
LEP students only				
Identified	339	2	472	3
Excluded	184	0	145	1
Assessed	155	1	327	2

0 Percentage is between 0.0 and 0.5.
SD = Students with Disabilities.
LEP = Limited English Proficient students.
NOTE: Within each grade level, the combined SD/LEP portion of the table is not a sum of the separate SD and LEP portions because some students were identified as both SD and LEP. Such students would be counted separately in the bottom portions, but counted only once in the top portion.
Within each portion of the table, percentages may not sum properly due to rounding..
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U. S. History Assessments.

BEST COPY AVAILABLE

Table A.6 Students Identified as SD and/or LEP Where Accommodations Were Permitted

Percentage of students identified as SD and/or LEP where accommodations were permitted, grades 4, 8, and 12: 2001

		Number of students sampled	Weighted percentage of students
Grade 4			
SD and/or LEP students	Identified	1,126	18
	Excluded	263	3
	Assessed	863	14
Assessed without accommodations		504	6
Assessed with accommodations		359	8
SD students only	Identified	645	13
	Excluded	142	2
	Assessed	503	11
Assessed without accommodations		180	3
Assessed with accommodations		323	7
LEP students only	Identified	584	6
	Excluded	154	1
	Assessed	430	4
Assessed without accommodations		336	3
Assessed with accommodations		94	1
Grade 8			
SD and/or LEP students	Identified	1,916	17
	Excluded	484	3
	Assessed	1,432	13
Assessed without accommodations		863	7
Assessed with accommodations		569	6
SD students only	Identified	1,308	13
	Excluded	312	2
	Assessed	996	10
Assessed without accommodations		438	4
Assessed with accommodations		558	6
LEP students only	Identified	715	4
	Excluded	201	1
	Assessed	514	3
Assessed without accommodations		445	3
Assessed with accommodations		69	#
Grade 12			
SD and/or LEP students	Identified	1,216	10
	Excluded	397	2
	Assessed	819	7
Assessed without accommodations		566	5
Assessed with accommodations		253	3
SD students only	Identified	834	8
	Excluded	327	2
	Assessed	507	5
Assessed without accommodations		276	3
Assessed with accommodations		231	2
LEP students only	Identified	417	2
	Excluded	82	#
	Assessed	335	2
Assessed without accommodations		301	2
Assessed with accommodations		34	#

Percentage is between 0.0 and 0.5.

SD = Students with Disabilities.

LEP = Limited English Proficient students.

NOTE: Within each grade level, the combined SD/LEP portion of the table is not a sum of the separate SD and LEP portions because some students were identified as both SD and LEP. Such students would be counted separately in the bottom portions but counted only once in the top portion.

Within each portion of the table, percentages may not sum properly due to rounding.

SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

BEST COPY AVAILABLE

Investigating the Effects of Exclusion Rates on Assessment Results

As indicated by the data in the previous section, exclusion rates have tended to increase across assessment years in the samples that did not permit accommodations. In considering the effects of exclusion rates on assessment results, at least one major issue becomes evident. If exclusion rates vary substantially across assessment years, then the ability to report trends (i.e., compare results between years) may be threatened by the fact that the results from different years are based on different proportions of the population.

NCES has funded research into ways in which excluded students might be included in the estimation of scores for total populations and has also commissioned studies of the impact of assessment accommodations on overall scores. Several statistical adjustment approaches for estimating full populations (including estimates for excluded students) have been proposed, but none has yet been judged ready for operational use. Regarding the impact of assessment accommodations on overall scores, ETS has conducted differential item functioning (DIF) studies of items assessed with accommodations in the 1996 assessment.[10] In these studies, ETS researchers found little evidence that accommodations changed the functioning of test questions.

Types of Accommodations Permitted

Table A.7 displays the number and the percentages of SD and/or LEP students assessed with the variety of available accommodations. It should be noted that students assessed with accommodations typically received some combination of accommodations. The numbers and percentages presented in the table reflect only the primary accommodation provided. For example, students assessed in small groups (as compared to standard NAEP sessions of about 30 students) usually received extended time. In one-on-one administrations, students often received assistance in recording answers and were afforded extra time. Extended time was considered the primary accommodation only when it was the sole accommodation provided.

[10] For information on DIF studies of items assessed with accommodations in the 1996 mathematics and science assessments, see Mazzeo, J. M., Carlson, J. E., Voelkl, K. E., & Lutkus, A. D. (1999). *Increasing the participation of special needs students in NAEP: A report on 1996 NAEP research activities.* Washington, DC: National Center for Education Statistics.

Table A.7 Students Identified as SD and/or LEP by Type of Accommodation

Percentage of students identified as SD and/or LEP by type of accommodation where accommodations were permitted, grades 4, 8, and 12: 2001

	Grade 4		Grade 8		Grade 12	
	Number of students sampled	Weighted percentage of students	Number of students sampled	Weighted percentage of students	Number of students sampled	Weighted percentage of students
SD and/or LEP students						
Bilingual dictionary	38	0.39	10	0.05	23	0.11
Large-print book	2	0.04	2	0.01	3	0.02
Extended time	30	0.38	104	1.09	86	0.72
Read aloud	26	0.65	15	0.11	13	0.15
Small group	238	5.80	416	4.81	124	1.55
One-on-one	19	0.38	7	0.10	0	0.00
Scribe/computer	5	0.17	2	0.08	2	0.02
Other	1	0.02	13	0.23	2	0.02
SD students only						
Bilingual dictionary	2	0.02	1	0.01	1	0.00
Large-print book	2	0.04	2	0.01	3	0.02
Extended time	30	0.38	104	1.09	86	0.72
Read aloud	26	0.65	15	0.11	13	0.15
Small group	238	5.80	416	4.81	124	1.55
One-on-one	19	0.38	7	0.10	0	0.00
Scribe/computer	5	0.17	2	0.08	2	0.02
Other	1	0.02	11	0.18	2	0.02
LEP students only						
Bilingual dictionary	38	0.39	10	0.05	23	0.11
Large-print book	0	0.00	0	0.00	0	0.00
Extended time	15	0.15	25	0.16	8	0.04
Read aloud	6	0.05	1	0.01	0	0.00
Small group	30	0.25	31	0.19	3	0.02
One-on-one	4	0.04	0	0.00	0	0.00
Scribe/computer	0	0.00	0	0.00	0	0.00
Other	1	0.02	2	0.05	0	0.00

SD = Students with Disabilities. LEP = Limited English Proficient students.
NOTE: The combined SD/LEP portion of the table is not a sum of the separate SD and LEP portions because some students were identified as both SD and LEP. Such students would be counted separately in the bottom portions, but counted only once in the top portion.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U. S. History Assessment.

BEST COPY AVAILABLE

Data Collection and Scoring

The 2001 U.S. history assessment was conducted from January through March 2001, with some makeup sessions in early April. As with all NAEP assessments, data collection for the 2001 assessment was conducted by a trained field staff. This was accomplished by staff from Westat, Inc.

Materials from the 2001 assessment were shipped to NCS Pearson, where trained staff evaluated the responses to the constructed-response questions using scoring rubrics or guides prepared by ETS. Each constructed-response question had a unique scoring rubric that defined the criteria used to evaluate students' responses. The extended constructed-response questions were evaluated with four-level rubrics, and almost all of the short constructed-response questions were rated according to three-level rubrics that permitted partial credit. Other short constructed-response questions were scored as either acceptable or unacceptable.

For the 2001 U.S. history assessment, approximately 315,000 constructed responses were scored. This number includes rescoring to monitor inter-rater reliability. The within-year average percentage of agreement for the 2001 national reliability sample was 93 percent at grade 4, 91 percent at grade 8, and 88 percent at grade 12.

Data Analysis and IRT Scaling

Subsequent to the professional scoring, all information was transcribed to the NAEP database at ETS. Each processing activity was conducted with rigorous quality control. After the assessment information was compiled in the database, the data were weighted according to the population structure. The weighting for the national sample reflected the probability of selection for each student as a result of the sampling design, adjusted for nonresponse. Through post-stratification, the weighting assured that the representation of certain subpopulations corresponded to figures from the U.S. Census and the Current Population Survey.[11]

Analyses were then conducted to determine the percentages of students who gave various responses to each cognitive and background question. In determining these percentages for the cognitive questions, a distinction was made between missing responses at the end of a block (i.e., missing responses subsequent to the last question the student answered) and missing responses prior to the last observed response. Missing responses before the last observed response were considered intentional omissions. In analysis, omitted responses to multiple-choice items were scored as fractionally correct.[12] For constructed-response items, omitted responses were placed into the lowest score category.

---

[11] These procedures are described more fully in the "Weighting and Variance Estimation" section later in this document. For additional information about the use of weighting procedures, see the forthcoming *NAEP 2001 Technical Report*. In addition, the reader may consult the *NAEP 2000 Technical Report* for a discussion of weighting procedures that are common to all NAEP assessments.

[12] Lord, F. M. (1980). *Applications of item response theory to practical testing problems.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Missing responses at the end of the block were considered "not reached" and treated as if the questions had not been presented to the student. In calculating response percentages for each question, only students classified as having been presented the question were included in the denominator of the statistic.

It is standard NAEP practice to treat all nonrespondents to the last question in a block as if they had not reached the question. For multiple-choice and short constructed-response questions, this practice produces a reasonable pattern of results in that the proportion reaching the last question is not dramatically smaller than the proportion reaching the next-to-last question. However, for history blocks that ended with extended constructed-response questions, the standard practice would result in extremely large drops in the proportion of students attempting the final question. Therefore, for blocks ending with an extended constructed-response question, students who answered the next-to-last question but did not respond to the extended constructed-response question were classified as having intentionally omitted the last question.

Item Response Theory (IRT) was used to estimate average history scale scores for the nation and for various subgroups of interest within the nation. IRT models the probability of answering a question in a certain way as a mathematical function of proficiency or skill. The main purpose of IRT analysis is to provide a common scale on which performance can be compared across groups such as those defined by characteristics, including gender and race/ethnicity.

In producing the U.S. history scales, three distinct IRT models were used. Multiple-choice questions were scaled using the three-parameter logistic (3PL) model; short constructed-response questions rated as acceptable or unacceptable were scaled using the two-parameter logistic (2PL) model; and short constructed-response questions rated according to a three-level rubric, as well as extended constructed-response questions rated on a four-level rubric, were scaled using a Generalized Partial-Credit (GPC) model.[13] Developed by ETS and first used in 1992, the GPC model permits the scaling of questions scored according to multipoint rating schemes. The model takes full advantage of the information available

[13] Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement, (16)2*, 159–176.

from each of the student response categories used for these more complex constructed-response questions.[14]

The U.S. history scale is composed of three types of questions: multiple-choice, short constructed-response (scored either dichotomously or allowing for partial credit), and extended constructed-response (scored according to a partial-credit model). One question about the U.S. history scales concerns the amount of information contributed by each type of question. Unfortunately, this question has no simple answer for the NAEP U.S. history assessment, due to the procedures used to form the composite history scale. The information provided by a given question is determined by the IRT model used to scale the question. It is a function of the item parameters and varies by level of U.S. history proficiency.[15] Thus, the answer to the query "How much information do the different types of questions provide?" will differ for each level of U.S. history performance. When considering the composite U.S. history scale, the answer is even more complicated. The U.S. history data are scaled separately by the four themes (change and continuity in American democracy: ideas, institutions, practices, and controversies; the gathering and interactions of peoples, cultures, and ideas; economic and technological changes and their relation to society, ideas, and the environment; and the changing role of America in the world), resulting in four separate subscales at each grade. The composite scale is a weighted combination of these subscales. IRT information functions are only strictly comparable when the item parameters are estimated together. Because the composite scale is based on four separate estimation runs, there is no direct way to compare the information provided by the questions on the composite scale.

Because of the BIB-spiraling design used by NAEP, students do not receive enough questions about a specific topic to provide reliable information about individual performance. (For more information on BIB-spiraling, see "The Assessment Design" section presented earlier in this appendix.) Traditional test scores for individual students, even those based on IRT, would lead to misleading estimates of population characteristics, such as subgroup means and percentages of students at or above a certain scale-score level. Consequently, NAEP constructs sets of plausible values designed to represent the distribution of performance in the population. A plausible value for an individual is not a scale score for that individual, but may be regarded as a representative value from the distribution of potential scale scores for all students in the population with similar characteristics and identical patterns of item response. Statistics describing performance on the NAEP U.S. history scale are based on the plausible

[14] More detailed information regarding the IRT analyses used in NAEP assessments will be provided in the forthcoming *NAEP 2001 Technical Report*. In addition, the reader may consult the *NAEP 2000 Technical Report* for a discussion of analysis procedures that are common to all NAEP assessments.

[15] Donoghue, J. R. (1994). An empirical examination of the IRT information of polytomously scored reading items under the generalized partial credit model. *Journal of Educational Measurement, (31)*4, 295–311.

values. Under the assumptions of the scaling models, these population estimates will be consistent, in the sense that the estimates approach the model-based population values as the sample size increases, which would not be the case for population estimates obtained by aggregating optimal estimates of individual performance.[16]

Item Mapping Procedures

The U.S. history performance of fourth-, eighth-, and twelfth-graders can be illustrated by "item maps," which position question or "item" descriptions along the NAEP U.S. history scale at each grade. Each question shown is placed at the point on the scale where questions are likely to be answered successfully by students. The descriptions used on these maps focus on the U.S. history knowledge or skill needed to answer the question. For multiple-choice questions, the description indicates the knowledge or skill demonstrated by selection of the correct option; for constructed-response questions, the description takes into account the knowledge or skill specified by the different levels of scoring criteria for that question.

To map questions to particular points on the NAEP U.S. history scale, a response probability convention was adopted that would divide those who had a higher probability of success from those who had a lower probability. Establishing a response probability convention has an impact on the mapping of the test questions onto the U.S. history scale. A lower boundary convention maps the history questions at lower points along the scale, and a higher boundary convention maps the same questions at higher points on the scale. The underlying distribution of U.S. history skills in the population does not change, but the choice of a response probability convention does have an impact on the proportion of the student population that is reported as "able to do" the questions on the U.S. history scales.

There is no obvious choice of a point along the probability scale that is clearly superior to any other point. If the convention were set with a boundary at 50 percent, those above the boundary would be more likely to get a question right than get it wrong, while those below the boundary would be more likely to get the question wrong than right. Although this convention has some intuitive appeal, it was rejected on the grounds that having a 50/50 chance of getting the question right shows an insufficient degree of mastery. If the convention were set with a boundary at 80 percent, students above the criterion would have a high probability of success with a question. However, many students below this criterion show some level of U.S. history ability that would be ignored by such a stringent criterion. In particular, those in the range between 50 and 80 percent correct would be more likely to get the question right than wrong, yet would not be in the group described as "able to do" the question.

In a compromise between the 50 percent and the 80 percent conventions,

[16] For theoretical and empirical justification of the procedures employed, see Mislevy, R. J. (1988). Randomization-based inferences about latent variables from complex samples. *Psychometrika, (56)*2, 177–196.

For computational details, see the forthcoming *NAEP 2001 Technical Report*.

NAEP has adopted two related response probability conventions: 65 percent for constructed-response questions (where guessing is not a factor) and 74 percent for multiple-choice questions with four response options (to correct for the possibility of answering correctly by guessing). These probability conventions were established, in part, based on an intuitive judgment that they would provide the best picture of students' U.S. history skills.

Some additional support for the dual conventions adopted by NAEP was provided by Huynh.[17] He examined the IRT information provided by items, according to the IRT model used in scaling NAEP questions. ("Information" is used here in a technical sense. See the forthcoming *NAEP 2001 Technical Report* for details.) Following Bock, Huynh decomposed the item information into that provided by a correct response [P(q) I(q)] and that provided by an incorrect response [(1– P(q)) I(q)].[18] Huynh showed that the item information provided by a correct response to a constructed-response item is maximized at the point along the U.S. history scale at which the probability of a correct response is 0.65 (for multiple-choice items, the information provided by a correct response is maximized at the point at which the probability of getting the item correct is 0.74). It should be noted, however, that maximizing the item information I(q), rather than the information provided by a correct response [P(q) I(q)], would imply an item mapping criterion closer to 50 percent.

Results are presented in terms of the composite U.S. history scale. However, the U.S. history assessment was scaled separately for the four themes in history at grades 4, 8, and 12. The composite scale is a weighted combination of the four subscales for the four themes in U.S. history. To obtain item map information, a procedure developed by Donoghue was used.[19] This method models the relationship between the item response function for the subscale and the subscale structure to derive the relationship between the item score and the composite scale (i.e., an item response function for the composite scale). This item response function is then used to derive the probability used in the mapping.

Weighting and Variance Estimation

A multistage sampling design was used to select the students who were assessed. The properties of a sample selected through such a design could be very different from those of a simple random sample, in which every student in the target population has an equal chance of selection and in which the observations from different sampled students can be considered to be statistically independent of one another. Therefore, the properties of the sample for the data collection design were taken into account during the analysis of the assessment data.

One way that the properties of the sample design were addressed was by using sampling weights to account for the fact that the probabilities of selection were not

[17] Huynh, H. (1994, October). *Some technical aspects of standard setting*. Paper presented at the Joint Conference on Standard Setting for Large-Scale Assessment, Washington, DC.

[18] Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more latent categories. *Psychometrika*, 37, 29–51.

[19] Donoghue, J. R. (1997, March). *Item mapping to a weighted composite scale*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

identical for all students. All population and subpopulation characteristics based on the assessment data were estimated using sampling weights. These weights included adjustments for school and student nonresponse.

Not only must appropriate estimates of population characteristics be derived, but appropriate measures of the degree of uncertainty must be obtained for those statistics. Two components of uncertainty are accounted for in the variability of statistics based on student ability: 1) the uncertainty due to sampling only a relatively small number of students, and 2) the uncertainty due to sampling only a relatively small number of cognitive questions. The first component accounts for the variability associated with the estimated percentages of students who had certain background characteristics or who answered a certain cognitive question correctly.

Because NAEP uses multistage sampling procedures, conventional formulas for estimating sampling variability that assume simple random sampling are inappropriate. NAEP uses a jackknife replication procedure to estimate standard errors. The jackknife standard error provides a reasonable measure of uncertainty for any student information that can be observed without error. However, because each student typically responds to only a few questions within any theme of history, the scale score for any single student would be imprecise. In this case, plausible values methodology can be used to describe the performance of groups and subgroups of students, but the underlying imprecision involved in this step adds another component of variability to statistics based on NAEP scale scores.[20]

Typically, when the standard error is based on a small number of students or when the group of students is enrolled in a small number of schools, the amount of uncertainty associated with the estimation of standard errors may be quite large. Estimates of standard errors subject to a large degree of uncertainty are followed by the "!" symbol to indicate that the nature of the sample does not allow accurate determination of the variability of the statistic. In such cases, the standard errors—and any confidence intervals or significance tests involving these standard errors—should be interpreted cautiously. Additional details concerning procedures for identifying such standard errors are discussed in the forthcoming *NAEP 2001 Technical Report*.

Drawing Inferences from the Results

The reported statistics are estimates and are therefore subject to a measure of uncertainty. There are two sources of such uncertainty. First, NAEP uses a sample of students rather than testing all students. Second, all assessments have some amount of uncertainty related to the fact that they cannot ask all questions that might be

[20] For further details, see Johnson, E. G. & Rust, K. F. (1992). Population inferences and variance estimation for NAEP data. *Journal of Educational Statistics, (17)2*, 175–190.

asked in a content area. The magnitude of this uncertainty is reflected in the standard error of each of the estimates. When the percentages or average scale scores of certain groups are compared, the standard error should be taken into account, and observed similarities or differences should not be relied on solely. Therefore, the comparisons are based on statistical tests that consider the standard errors of those statistics and the magnitude of the difference among the averages or percentages.

Using confidence intervals based on the standard errors provides a way to take into account the uncertainty associated with sample estimates and to make inferences about the population averages and percentages in a manner that reflects that uncertainty. An estimated sample average scale score plus or minus 1.96 standard errors approximates a 95 percent confidence interval for the corresponding population quantity. This statement means that one can conclude with approximately a 95 percent level of confidence that the average performance of the entire population of interest (e.g., all fourth-grade students in public and nonpublic schools) is within plus or minus 1.96 standard errors of the sample average.

As an example, suppose that the average U.S. history scale score of the students in a particular group was 256 with a standard error of 1.2. An approximate 95 percent confidence interval for the population quantity would be as follows:

$$\text{Average} \pm 1.96 \text{ standard errors}$$
$$256 \pm 1.96 \times 1.2$$
$$256 \pm 2.35$$
$$(253.65, 258.35)$$

Thus, one can conclude with a 95 percent level of confidence that the average scale score for the entire population of students in that group is between 253.65 and 258.35. It should be noted that this example, and the examples in the following sections, are illustrative. More precise estimates carried out to one or more decimal places are used in the actual analyses.

Similar confidence intervals can be constructed for percentages, if the percentages are not extremely large or extremely small. Extreme percentages should be interpreted with caution. Adding or subtracting the standard errors associated with extreme percentages could cause the confidence interval to exceed 100 percent or go below 0 percent, resulting in numbers that are not meaningful. The forthcoming *NAEP 2001 Technical Report* will contain a more complete discussion of extreme percentages.

Analyzing Group Differences in Averages and Percentages

Statistical tests determine whether the evidence, based on the data from the groups in the sample, is strong enough to conclude that the averages or percentages are actually different for those groups in the population. If the evidence is strong (i.e., the difference is statistically significant), the report describes the group averages or percentages as being different (e.g., one group performed higher than or lower than another group), regardless of whether the sample averages or percentages appear to be approximately the same.

The reader is cautioned to rely on the results of the statistical tests rather than on the apparent magnitude of the difference between sample averages or percentages when determining whether the sample differences are likely to represent actual differences among the groups in the population.

To determine whether a real difference exists between the average scale scores (or percentages of a certain attribute) for two groups in the population, one needs to obtain an estimate of the degree of uncertainty associated with the difference between the averages (or percentages) of these groups for the sample. This estimate of the degree of uncertainty, called the "standard error of the difference" between the groups, is obtained by taking the square of each group's standard error, summing the squared standard errors, and taking the square root of that sum.

$$\text{Standard Error of the Difference} = SE_{A-B} = \sqrt{(SE_A^2 + SE_B^2)}$$

Similar to how the standard error for an individual group average or percentage is used, the standard error of the difference can be used to help determine whether differences among groups in the population are real. The difference between the averages or percentages of the two groups plus or minus two standard errors of the difference represents an approximate 95 percent confidence interval. If the resulting interval includes zero, there is insufficient evidence to claim a real difference between the groups in the population. If the interval does not contain zero, the difference between the groups is statistically significant (different) at the 0.05 level.

As an example of comparing groups, consider the problem of determining whether the average U.S. history scale score of group A is higher than that of group B. Suppose that the sample estimates of the average scale scores and standard errors were as follows:

Group	Average Scale Score	Standard Error
A	218	0.9
B	216	1.1

The difference between the estimates of the average scale scores of groups A and B is two points (218 – 216). The standard error of this difference is

$$\sqrt{(0.9^2 + 1.1^2)} = 1.4$$

Thus, an approximate 95 percent confidence interval for this difference is plus or minus two standard errors of the difference

$$2 \pm 1.96 \times 1.4$$
$$2 \pm 2.74$$
$$(-0.74, 4.74)$$

The value zero is within the confidence interval; therefore, there is insufficient evidence to claim that group A outperformed group B.

Conducting Multiple Tests

The procedures in the previous section and the certainty ascribed to intervals (e.g., a 95 percent confidence interval) are based on statistical theory that assumes that only one confidence interval or test of statistical significance is being performed. However, many different groups are being compared (i.e., multiple sets of confidence intervals are being analyzed). In sets of confidence intervals, statistical theory indicates that the certainty associated with the entire set of intervals is less than that attributable to each individual comparison from the set. To hold the significance level for the set of comparisons at a particular level (e.g., 0.05), adjustments (called "multiple comparison procedures"[21]) must be made to the methods described in the previous section. One such procedure, the False Discovery Rate (FDR) procedure,[22] was used to control the certainty level.

Unlike the other multiple comparison procedures (e.g., the Bonferroni procedure) that control the familywise error rate (i.e., the probability of making even one false rejection in the set of comparisons), the FDR procedure controls the expected proportion of falsely rejected hypotheses. Furthermore, familywise procedures are considered conservative for large families of comparisons.[23] Therefore, the FDR procedure is more suitable for multiple comparisons in NAEP than other procedures. A detailed description of the FDR procedure appears in the forthcoming *NAEP 2001 Technical Report*.

To illustrate how the FDR procedure is used, consider the comparisons of current and previous years' average U.S. history scale scores for the five groups presented in table A.8. Note that the difference in average scale scores and the standard error of the difference are calculated in a way comparable with that of the example in the previous section. The test statistic shown is the difference in average scale scores divided by the standard error of the difference.

The difference in average scale scores and its standard error can be used to find an approximate 95 percent confidence interval as in the example in the previous section or they can be used to identify a

---

21 Miller, R. G. (1966). *Simultaneous statistical inference*. New York: Wiley.

22 Benjamini, Y. & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B, No. 1.*, pp 289–300.

23 Williams, V. S. L., Jones, L. V., & Tukey, J. W. (1999). *Controlling error in multiple comparisons with examples from state-to-state differences in educational achievement*. Journal of Educational and Behavioral Statistics, *24*(1), 42–69.

Table A.8 FDR Comparisons of Average Scale Scores

Example of FDR comparisons of average scale scores for different groups of students

	Previous year		Current year		Previous year and current year			
	Average scale score	Standard error	Average scale score	Standard error	Difference in averages	Standard error of difference	Test statistic	Percent confidence*
Group 1	224	1.3	226	1.0	2.08	1.62	1.29	20
Group 2	187	1.7	193	1.7	6.31	2.36	2.68	1
Group 3	191	2.6	197	1.7	6.63	3.08	2.15	4
Group 4	229	4.4	232	4.6	3.24	6.35	.51	62
Group 5	201	3.4	196	4.7	–5.51	5.81	–.95	35

* The percent confidence is 2(1–F(x)) where F(x) is the cumulative distribution of the t-distribution with the degrees of freedom adjusted to reflect the complexities of the sample design.

confidence percentage. In the example in the previous section, because an approximate 95 percent confidence interval was desired, the number 1.96 was used to multiply the standard error of the difference to create the approximate confidence interval. In the current example, the confidence interval for the test statistics is identified from statistical tables. Instead of checking to see if zero is within the 95 percent confidence interval about the mean, the significance level from the statistical tables can be directly compared to 100–95 = 5 percent.

If the comparison of average scale scores across two years were made for only one of the five groups, there would be a significant difference between the average scale scores for the two years if the significance level were less than 5 percent. However, because we are interested in the difference in average scale scores across the two years for all five of the groups, comparing each of the significance levels to 5 percent is not adequate. Groups of students defined by shared characteristics, such as race/ethnicity groups, are treated as sets or families when making comparisons. However, comparisons of average scale scores for each pair of years were treated separately. So the steps described in this example would be replicated for the comparison of other current and previous year average scale scores.

To use the FDR procedure to take into account that all comparisons are of interest to us, the percents of confidence in the example are ordered from largest to smallest: 62, 35, 20, 4, and 1. In the FDR procedure, 62 percent confidence for the Group 4 comparison would be compared to 5 percent, 35 percent for the Group 5 comparison would be compared to 0.05*(5–1)/5 = 0.04*100 = 4 percent,[24] 20 percent for the Group 1 comparison

[24] The level of confidence times the number of comparisons minus one divided by the number of comparisons is 0.05*(5–1)/5 = 0.04*100 = 4 percent.

would be compared to 0.05*(5–2)/5 = 0.03*100 = 3 percent, 4 percent for the Group 3 comparison would be compared to 0.05*(5–3)/5 = 0.02*100 = 2 percent, and 1 percent for the Group 2 comparison (actually slightly smaller than 1 prior to rounding) would be compared to 0.05*(5–4)/5 = 0.01*100 = 1 percent. The last of these comparisons is the only one for which the percent confidence is smaller than the FDR procedure value. The difference in the current year and previous years' average scale scores for the Group 2 students is significant; for all of the other groups, average scale scores for current and previous year are not significantly different from one another. In practice, a very small number of counter-intuitive results occur when using the FDR procedures to examine between-year differences in subgroup results by jurisdiction. In those cases, results were not included in this report. NCES is continuing to evaluate the use of FDR and multiple-comparison procedures for future reporting.

NAEP Reporting Groups

Results are provided for groups of students defined by shared characteristics—region of the country, gender, race or ethnicity, school's type of location, eligibility for the free/reduced-price school lunch program, and type of school. Based on participation rate criteria, results are reported for subpopulations only when sufficient numbers of students and adequate school representation are present. The minimum requirement is at least 62 students in a particular subgroup from at least five primary sampling units (PSUs).[25] However, the data for all students, regardless of whether their subgroup was reported separately, were included in computing overall results. Definitions of the subpopulations are presented below.

Region

Results in NAEP are reported for four regions of the nation: Northeast, Southeast, Central, and West. Figure A.1 shows how states are subdivided into these NAEP regions. All 50 states and the District of Columbia are listed. Other jurisdictions, including territories and the two Department of Defense Educational Activities jurisdictions are not assigned to any region.

Gender

Results are reported separately for males and females.

Race/Ethnicity

The race/ethnicity variable is derived from two questions asked of students and from school records, and it is used for race/ethnicity subgroup comparisons. Two questions from the set of general student background questions were used to determine race/ethnicity:

If you are Hispanic, what is your Hispanic background?

- ❑ I am not Hispanic
- ❑ Mexican, Mexican American, or Chicano
- ❑ Puerto Rican
- ❑ Cuban
- ❑ Other Spanish or Hispanic background

[25] For the national assessment, a PSU is a selected geographic region (a county, group of counties, or metropolitan statistical area). Further details about the procedure for determining minimum sample size appear in the *NAEP 2000 Technical Report* and the forthcoming *NAEP 2001 Technical Report*.

Figure A.1
States by Region

States included in the four NAEP regions: 2001

Northeast	Southeast	Central	West
Connecticut	Alabama	Illinois	Alaska
Delaware	Arkansas	Indiana	Arizona
District of Columbia	Florida	Iowa	California
Maine	Georgia	Kansas	Colorado
Maryland	Kentucky	Michigan	Hawaii
Massachusetts	Louisiana	Minnesota	Idaho
New Hampshire	Mississippi	Missouri	Montana
New Jersey	North Carolina	Nebraska	Nevada
New York	South Carolina	North Dakota	New Mexico
Pennsylvania	Tennessee	Ohio	Oklahoma
Rhode Island	*Virginia	South Dakota	Oregon
Vermont	West Virginia	Wisconsin	Texas
*Virginia			Utah
			Washington
			Wyoming

*The part of Virginia that is included in the Northeast region is the Washington, DC metropolitan area; the remainder of the state is included in the Southeast region.

Students who responded to this question by filling in the second, third, fourth, or fifth oval were considered Hispanic. For students who filled in the first oval, did not respond to the question, or provided information that was illegible or could not be classified, responses to the following question were examined to determine their race/ethnicity:

Which best describes you?

- ❑ White (not Hispanic)
- ❑ Black (not Hispanic)
- ❑ Hispanic ("Hispanic" means someone who is Mexican, Mexican American, Chicano, Puerto Rican, Cuban, or other Spanish or Hispanic background.)
- ❑ Asian or Pacific Islander ("Asian or Pacific Islander" means someone who is from a Chinese, Japanese, Korean, Filipino, Vietnamese, Asian American or some other Asian or Pacific Islander background.)
- ❑ American Indian or Alaskan Native ("American Indian or Alaskan Native" means someone who is from one of the American Indian tribes or one of the original people of Alaska.)
- ❑ Other (specify) _____

Students' race/ethnicity was then assigned on the basis of their responses. For students who filled in the sixth oval ("Other"), provided illegible information or information that could not be classified, or did not respond at all, race/ethnicity was assigned as determined by school records.

Race/ethnicity could not be determined for students who did not respond to either of the demographic questions and whose schools did not provide information about race/ethnicity.

Also, some students indicated that they were from a Hispanic background (e.g., Puerto Rican or Cuban) and that a

BEST COPY AVAILABLE

racial/ethnic category other than Hispanic best described them. These students were classified as Hispanic based on the rules described above.

Type of Location

Results from the 2001 assessment are reported for students attending schools in three mutually exclusive location types: central city, urban fringe/large town, and rural/small town:

Central City: This category includes central cities of all Standard Metropolitan Statistical Areas (SMSA) as defined by the Office of Management and Budget. Central City is a geographical term and is not synonymous with "inner city."

Urban Fringe/Large Town: The urban fringe category includes all densely settled places and areas within SMSA's that are classified as urban by the Bureau of the Census, but which do not qualify as Central City. A Large Town is defined as a place outside a SMSA with a population greater than or equal to 25,000.

Rural/Small Town: Rural includes all places and areas with populations of less than 2,500 that are classified as rural by the Bureau of the Census. A Small Town is defined as a place outside a SMSA with a population of less than 25,000, but greater than or equal to 2,500.

Results for each type of location are not compared across years. This was due to new methods used by NCES to identify the type of location assigned to each school in the Common Core of Data (CCD). The new methods were put into place by NCES in order to improve the quality of the assignments and they take into account more information about the exact physical location of the school. The variable was revised in NAEP beginning with the 2000 assessments.

Eligibility for the Free/Reduced-Price School Lunch Program

Based on available school records, students were classified as either currently eligible for the free/reduced-price school lunch component of the Department of Agriculture's National School Lunch Program or not eligible. Eligibility for the program is determined by students' family income in relation to the federally established poverty level. Free lunch qualification is set at 130 percent of the poverty level, and reduced-price lunch qualification is set at 170 percent of the poverty level. The classification applies only to the school year when the assessment was administered (i.e., the 2000–2001 school year) and is not based on eligibility in previous years. If school records were not available, the student was classified as "Information not available." If the school did not participate in the program, all students in that school were classified as "Information not available."

Type of School

Results are reported by the type of school that the student attends—public or nonpublic. Nonpublic schools include Catholic and other private schools.[26] Because they are funded by federal authorities, not state/local governments, Bureau of Indian Affairs (BIA) schools and Department of Defense

[26] Through a pilot study, more detailed breakdowns of nonpublic school results are available on the NAEP Web Site (http://nces.ed.gov/nationsreportcard/history/results/index.asp).

Domestic Dependent Elementary and Secondary Schools (DDESS) are not included in either the public or nonpublic categories; they are included in the overall national results.

Grade 12 Participation Rates

NAEP has been described as a "low-stakes" assessment. That is, students receive no individual scores, and their NAEP performance has no effect on their grades, promotions, or graduation. There has been continued concern that this lack of consequences affects participation rates of students and schools, as well as the motivation of students to perform well on NAEP. Of particular concern has been the performance of twelfth-graders, who typically have lower student participation rates than fourth- and eighth-graders, and who are more likely to omit responses compared to the younger cohorts.

In NAEP, there has been a consistent pattern of lower participation rates for older students. In the 2001 NAEP assessments, for example, the student participation rates were 96 percent and 93 percent at grades 4 and 8, respectively. At grade 12, however, the participation rate was 78 percent. School participation rates (the percentage of sampled schools that participated in the assessment) have also typically decreased with grade level. Again citing the 2001 assessments, the school participation rate was 88 percent for the fourth grade, 87 percent for the eighth grade, and 80 percent for the twelfth grade.

The effect of participation rates on student performance, however, is unclear. Students may choose not to participate in NAEP for many reasons, such as desire to attend regular classes so as not to miss important instruction or conflict with other school-based activities. Similarly, there are a variety of reasons for which various schools do not participate. The sampling weights and nonresponse adjustments, described earlier in this document, provide an approximate statistical adjustment for nonparticipation. However, the effect of some school and student nonparticipation may have some undetermined effect on results.

More research is needed to delineate the factors that contribute to nonparticipation and lack of motivation. To that end, NCES is currently investigating how various types of incentives can be effectively used to increase participation in NAEP. One report that examines the impact of monetary incentives on student effort and performance is available on the NCES Web Site at http://nces.ed.gov/pubsearch/. Enter NCES#: 2001024.

Cautions in Interpretations

As described earlier, the NAEP U.S. history scale makes it possible to examine relationships between students' performance and various background factors measured by NAEP. However, a relationship that exists between achievement and another variable does not reveal its underlying cause, which may be influenced by a number of other variables. Similarly, the assessments do not capture the influence of unmeasured variables. The results are most useful when they are considered in combination with other knowledge about the student population and the educational system, such as trends in instruction, changes in the school-age population, and societal demands and expectations.

Appendix B
Data Appendix

This appendix contains complete data for all the tables and figures presented in this report, including average scores, achievement-level results, and percentages of students. In addition, standard errors appear in parentheses next to each scale score and percentage. The comparisons presented in this report are based on statistical tests that consider the magnitude of the difference between group averages or percentages and the standard errors of those statistics. Because NAEP scores and percentages are based on samples rather than the entire population(s), the results are subject to a measure of uncertainty reflected in the standard errors of the estimates. It can be said with 95 percent certainty that for each population of interest, the value for the whole population is within plus or minus two standard errors of the estimate for the sample. As with the figures and tables in the chapters, significant differences between results of previous assessments and the 2001 assessment are highlighted.

Appendix Focus

Complete data for all tables and figures.

Appendix Contents

Average Scores

Achievement-Level Results

Percentages of Students

Standard Errors

Table B.1: Data for Figure 2.1 National Scale Score Results

Average U.S. history scale scores, grades 4, 8, and 12: 1994 and 2001

	Grade 4	Grade 8	Grade 12
1994	205 (1.0)	259 (0.6)	286 (0.8)
2001	209 (1.0) *	262 (0.8) *	287 (1.0)

Standard errors of the estimated scale scores appear in parentheses.
* Significantly different from 1994.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

Table B.2: Data for Figure 2.2: National Performance Distribution

National U.S. history scale score percentiles, grades 4, 8, and 12: 1994 and 2001

		10th	25th	50th	75th	90th
Grade 4	1994	147 (2.1)	180 (1.5)	210 (0.9)	234 (1.2)	253 (1.4)
	2001	158 (1.6) *	186 (1.9) *	212 (1.3)	235 (0.9)	255 (1.4)
Grade 8	1994	217 (1.1)	239 (0.9)	261 (1.1)	282 (0.7)	299 (0.6)
	2001	220 (1.3)	241 (0.9) *	264 (0.8)	285 (0.7) *	303 (1.1) *
Grade 12	1994	243 (1.2)	265 (1.2)	288 (0.8)	309 (0.9)	326 (1.0)
	2001	246 (1.1)	266 (0.9)	288 (1.0)	309 (1.2)	327 (1.5)

Standard errors of the estimated scale scores appear in parentheses.
* Significantly different from 1994.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

BEST COPY AVAILABLE

Table B.3: Data for Figure 2.3: National Achievement-Level Results

Percentage of students within and at or above U.S. history achievement levels, grades 4, 8, and 12: 1994 and 2001

		Below *Basic*	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
Grade 4	1994	36 (1.1)	47 (0.9)	15 (0.9)	2 (0.3)	64 (1.1)	17 (1.0)
	2001	33 (1.1) *	49 (1.1)	16 (0.9)	2 (0.5)	67 (1.1) *	18 (1.0)
Grade 8	1994	39 (0.9)	48 (0.8)	13 (0.6)	1 (0.1)	61 (0.9)	14 (0.6)
	2001	36 (0.9) *	48 (0.9)	15 (0.8) *	2 (0.3) *	64 (0.9) *	17 (0.8) *
Grade 12	1994	57 (1.1)	32 (0.9)	10 (0.6)	1 (0.2)	43 (1.1)	11 (0.7)
	2001	57 (1.2)	32 (0.9)	10 (0.6)	1 (0.4)	43 (1.2)	11 (0.9)

Standard errors of the estimated percentages appear in parentheses.
* Significantly different from 1994.
NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

BEST COPY AVAILABLE

Table B.4. Data for Figure 3.1 National Scale Score Results by Gender

Percentage of students and average U.S. history scale scores by gender, grades 4, 8, and 12: 1994 and 2001

		Male	Female
Grade 4	1994	50 (0.8) 203 (1.5)	50 (0.8) 206 (1.1)
	2001	50 (0.8) 209 (1.1) *	50 (0.8) 209 (1.2) *
Grade 8	1994	50 (0.5) 259 (0.8)	50 (0.5) 259 (0.7)
	2001	49 (0.6) 264 (0.9) *	51 (0.6) 261 (0.9)
Grade 12	1994	50 (0.8) 288 (0.8)	50 (0.8) 285 (0.9)
	2001	49 (0.6) 288 (1.3)	51 (0.6) 286 (0.9)

The percentage of students is listed first with the corresponding average scale score presented below.
Standard errors of the estimated percentages and scale scores appear in parentheses.
* Significantly different from 1994.
NOTE: Percentages may not add to 100 due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

Table B.5. Data for Figure 3.2 National Scale Score Differences by Gender

Differences in average U.S. history scale scores by gender, grades 4, 8, and 12: 1994 and 2001

		Male-Female
Grade 4	1994	–2 (1.8)
	2001	–1 (1.6)
Grade 8	1994	# (1.0)
	2001	2 (1.2)
Grade 12	1994	3 (1.2)
	2001	2 (1.5)

Standard errors of the estimated difference in scale scores appear in parentheses.
Score differences are calculated based on differences between unrounded average scale scores.
Difference is between –0.5 and 0.5.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

BEST COPY AVAILABLE

Table B.6: Data for Figure 3.3 National Achievement-Level Results by Gender

Percentage of students within and at or above U.S. history achievement levels by gender, grades 4, 8, and 12: 1994 and 2001

			Below *Basic*	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
Grade 4	Male	1994	38 (1.6)	44 (1.5)	16 (1.3)	2 (0.4)	62 (1.6)	18 (1.4)
		2001	34 (1.3)	47 (1.3)	17 (1.1)	2 (0.7)	66 (1.3)	19 (1.2)
	Female	1994	35 (1.4)	50 (1.5)	14 (1.0)	2 (0.4)	65 (1.4)	16 (1.1)
		2001	32 (1.4)	51 (1.6)	15 (1.1)	2 (0.4)	68 (1.4)	17 (1.1)
Grade 8	Male	1994	39 (1.0)	47 (0.9)	14 (0.9)	1 (0.3)	61 (1.0)	15 (0.8)
		2001	35 (1.1) *	47 (1.3)	17 (1.0)	2 (0.3) *	65 (1.1) *	18 (1.0) *
	Female	1994	39 (1.3)	49 (1.2)	12 (0.8)	1 (0.1)	61 (1.3)	13 (0.8)
		2001	37 (1.2)	48 (1.0)	14 (0.8)	1 (0.4)	63 (1.2)	15 (0.8)
Grade 12	Male	1994	55 (1.2)	34 (0.9)	11 (0.7)	1 (0.3)	45 (1.2)	12 (0.7)
		2001	55 (1.6)	33 (1.1)	11 (0.8)	1 (0.5)	45 (1.6)	12 (1.1)
	Female	1994	60 (1.4)	31 (1.3)	9 (0.7)	1 (0.2)	40 (1.4)	9 (0.8)
		2001	59 (1.3)	31 (1.2)	9 (0.7)	1 (0.3)	41 (1.3)	10 (0.9)

Standard errors of the estimated percentages appear in parentheses.
* Significantly different from 1994.
NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

BEST COPY AVAILABLE

Table B.7: Data for Figure 3.4 National Scale Score Results by Race/Ethnicity

Percentage of students and average U.S. history scale scores by race/ethnicity, grades 4, 8, and 12: 1994 and 2001

		White	Black	Hispanic	Asian/ Pacific Islander	American Indian
Grade 4	1994	69 (0.3) 215 (1.2)	15 (0.1) 177 (1.6)	11 (0.2) 180 (2.7)	3 (0.3) 205 (3.9)	2 (0.3) 190 (6.1)
	2001	65 (0.4) 220 (1.1) *	14 (0.2) 188 (1.8) *	16 (0.3) 186 (2.5)	3 (0.2) 213 (2.7)	2 (0.2) 197 (6.9)
Grade 8	1994	69 (0.2) 267 (0.8)	15 (0.1) 239 (1.4)	11 (0.1) 243 (1.3)	3 (0.4) 263 (4.0)	1 (0.3) 246 (3.7) !
	2001	67 (0.4) 271 (0.8) *	14 (0.2) 243 (1.8)	14 (0.2) 243 (1.5)	4 (0.2) 267 (3.4)	1 (0.2) 249 (4.5)
Grade 12	1994	74 (0.4) 292 (0.8)	12 (0.2) 265 (1.5)	9 (0.3) 267 (1.6)	4 (0.2) 285 (3.4)	1 (0.2) 279 (4.0) !
	2001	70 (0.4) 292 (1.0)	13 (0.2) 269 (1.5)	12 (0.2) 274 (1.7) *	5 (0.2) 295 (4.6)	1 (0.2) 277 (5.5) !

The percentage of students is listed first with the corresponding average scale score presented below.
Standard errors of the estimated percentages and scale scores appear in parentheses.
* Significantly different from 1994.
! The nature of the sample does not allow accurate determination of the variability of the statistic.
NOTE: Percentage may not add to 100 due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

Table B.8: Data for Figure 3.5 National Scale Score Differences by Race/Ethnicity

Differences in average U.S. history scale scores by race/ethnicity, grades 4, 8, and 12: 1994 and 2001

		White-Black	White-Hispanic
Grade 4	1994	38 (2.0)	35 (3.0)
	2001	31 (2.1) *	33 (2.7)
Grade 8	1994	28 (1.6)	24 (1.6)
	2001	28 (2.0)	28 (1.7)
Grade 12	1994	27 (1.7)	26 (1.8)
	2001	24 (1.8)	19 (2.0) *

Standard errors of the estimated difference in scale scores appear in parentheses.
*Significantly different from 1994.
Score differences are calculated based on differences between unrounded average scale scores.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

BEST COPY AVAILABLE

Table B.9: Data for Figure 3.6 National Achievement-Level Results by Race/Ethnicity

Percentage of students within and at or above U.S. history achievement levels by race/ethnicity, grades 4, 8, and 12: 1994 and 2001

			Below *Basic*	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
Grade 4	White	1994	26 (1.1)	52 (1.0)	19 (1.2)	3 (0.5)	74 (1.1)	22 (1.4)
		2001	21 (1.3) *	55 (1.4)	21 (1.4)	3 (0.7)	79 (1.3) *	24 (1.4)
	Black	1994	64 (1.8)	32 (2.0)	4 (1.0)	# (***)	36 (1.8)	4 (1.0)
		2001	56 (2.1) *	38 (1.9)	5 (0.9)	# (0.3)	44 (2.1) *	6 (1.0)
	Hispanic	1994	59 (3.6)	35 (3.3)	6 (1.2)	1 (***)	41 (3.6)	6 (1.2)
		2001	58 (3.0)	35 (2.6)	6 (1.1)	1 (0.3)	42 (3.0)	7 (1.1)
	Asian/Pacific Islander	1994	38 (3.4)	42 (4.6)	17 (4.2)	4 (1.9)	62 (3.4)	20 (3.8)
		2001	29 (3.8)	53 (4.6)	16 (2.7)	3 (1.9)	71 (3.8)	19 (3.2)
	American Indian	1994	49 (6.9)	41 (7.8)	9 (2.4)	# (***)	51 (6.9)	9 (2.7)
		2001	47 (6.4)	41 (6.0)	8 (3.0)	4 (***)	53 (6.4)	12 (4.6)
Grade 8	White	1994	29 (1.1)	54 (0.9)	16 (0.8)	1 (0.2)	71 (1.1)	17 (0.8)
		2001	25 (1.0)	53 (1.1)	19 (1.1)	2 (0.4) *	75 (1.0)	21 (1.1) *
	Black	1994	67 (2.3)	29 (2.2)	4 (0.6)	# (0.1)	33 (2.3)	4 (0.6)
		2001	62 (2.4)	34 (2.1)	4 (0.8)	# (***)	38 (2.4)	4 (0.8)
	Hispanic	1994	59 (2.3)	36 (2.2)	5 (0.8)	# (***)	41 (2.3)	5 (0.7)
		2001	60 (1.7)	34 (1.5)	5 (0.7)	# (0.2)	40 (1.7)	5 (0.7)
	Asian/Pacific Islander	1994	35 (5.5)	46 (3.7)	17 (2.7)	2 (0.9)	65 (5.5)	19 (3.0)
		2001	32 (3.8)	48 (3.0)	18 (3.2)	2 (0.8)	68 (3.8)	20 (3.6)
	American Indian	1994	58 (5.8) !	37 (5.9) !	5 (2.6) !	0 (***) !	42 (5.8) !	5 (2.6) !
		2001	50 (7.1)	42 (6.7)	7 (3.2)	1 (***)	50 (7.1)	8 (3.5)
Grade 12	White	1994	50 (1.2)	37 (0.9)	12 (0.8)	1 (0.2)	50 (1.2)	13 (0.8)
		2001	51 (1.4)	36 (1.1)	12 (0.8)	1 (0.4)	49 (1.4)	13 (1.0)
	Black	1994	83 (1.6)	15 (1.3)	2 (0.8)	# (***)	17 (1.6)	2 (0.8)
		2001	80 (1.5)	18 (1.4)	3 (0.6)	# (***)	20 (1.5)	3 (0.6)
	Hispanic	1994	78 (2.1)	18 (2.1)	4 (0.8)	# (***)	22 (2.1)	4 (0.7)
		2001	74 (2.4)	21 (1.8)	5 (1.0)	# (***)	26 (2.4)	5 (1.1)
	Asian/Pacific Islander	1994	57 (4.8)	29 (3.0)	12 (3.5)	2 (0.9)	43 (4.8)	13 (3.7)
		2001	47 (5.1)	31 (2.4)	17 (4.1)	5 (2.3)	53 (5.1)	21 (6.0)
	American Indian	1994	70 (7.6) !	25 (7.3) !	5 (2.3) !	0 (***) !	30 (7.6) !	5 (2.3) !
		2001	66 (7.2) !	33 (7.4) !	1 (***) !	0 (***) !	34 (7.2) !	1 (***) !

Standard errors of the estimated percentages appear in parentheses.
* Significantly different from 1994.
# Percentage is between 0.0 and 0.5.
! The nature of the sample does not allow accurate determination of the variability of the statistic.
(***) Standard error estimates cannot be accurately determined.
NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 History Assessments.

BEST COPY AVAILABLE

Table B.10: Data for Figure 3.7 National Scale Score Results by Region of the Country

Percentage of students and average U.S. history scale scores by region of the country, grades 4, 8, and 12: 1994 and 2001

		Northeast	Southeast	Central	West
Grade 4	1994	22 (0.7) 204 (2.4)	23 (1.0) 201 (1.9)	25 (0.8) 212 (2.6)	30 (0.6) 202 (2.1)
	2001	21 (0.8) 215 (2.5) *	24 (1.3) 208 (2.6)	24 (0.4) 217 (2.0)	31 (1.4) 200 (2.3)
Grade 8	1994	20 (0.8) 266 (1.7)	25 (0.9) 251 (1.3)	24 (0.6) 266 (1.3)	31 (0.8) 256 (1.1)
	2001	20 (0.9) 269 (1.9)	23 (1.1) 261 (2.0) *	25 (0.5) 267 (1.7)	32 (1.3) 255 (1.3)
Grade 12	1994	20 (0.5) 289 (1.9)	23 (0.8) 282 (1.4)	27 (0.7) 288 (1.4)	30 (0.7) 286 (1.6)
	2001	21 (0.9) 289 (3.4)	22 (1.2) 284 (1.7)	26 (0.6) 289 (1.4)	31 (1.5) 286 (1.6)

The percentage of students is listed first with the corresponding average scale score presented below.
Standard errors of the estimated percentages and scale scores appear in parentheses.
* Significantly different from 1994.
NOTE: Percentage may not add to 100 due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

Table B.11: Data for Figure 3.8 National Achievement-Level Results by Region of the Country

Percentage of students within and at or above U.S. history achievement levels by region of the country, grades 4, 8, and 12: 1994 and 2001

			Below *Basic*	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
Grade 4	Northeast	1994	37 (2.4)	45 (2.4)	16 (1.9)	3 (0.7)	63 (2.4)	18 (2.1)
		2001	27 (3.1) *	50 (2.9)	20 (2.4)	3 (1.1)	73 (3.1) *	23 (2.9) *
	Southeast	1994	39 (2.3)	46 (2.4)	13 (1.4)	2 (0.5)	61 (2.3)	15 (1.5)
		2001	34 (2.7) *	50 (1.8)	14 (1.5)	2 (1.2)	66 (2.7) *	16 (2.2)
	Central	1994	29 (2.8)	51 (2.6)	17 (2.0)	3 (1.0)	71 (2.8)	20 (2.6)
		2001	25 (2.3) *	52 (2.6)	20 (2.1) *	3 (1.1)	75 (2.3) *	24 (2.4) *
	West	1994	39 (2.5)	45 (2.4)	14 (1.9)	1 (0.8)	61 (2.5)	16 (1.8)
		2001	41 (2.5)	46 (1.9)	12 (1.1)	1 (0.4)	59 (2.5)	13 (1.2)
Grade 8	Northeast	1994	31 (2.0)	50 (1.2)	18 (1.6)	1 (0.3)	69 (2.0)	19 (1.7)
		2001	28 (2.2)	50 (1.7)	20 (2.0) *	2 (0.8)	72 (2.2)	22 (2.1) *
	Southeast	1994	49 (1.9)	42 (1.4)	8 (0.8)	# (0.2)	51 (1.9)	9 (0.8)
		2001	38 (2.3) *	46 (1.5) *	14 (1.1) *	2 (0.4)	62 (2.3) *	16 (1.3) *
	Central	1994	31 (2.3)	52 (2.0)	16 (1.3)	1 (0.4)	69 (2.3)	17 (1.2)
		2001	29 (2.2)	53 (1.5)	17 (1.5) *	2 (0.5)	71 (2.2)	19 (1.5) *
	West	1994	42 (1.3)	47 (1.4)	10 (1.1)	1 (0.2)	58 (1.3)	11 (1.2)
		2001	45 (1.7) *	43 (1.6) *	11 (1.1)	1 (0.2)	55 (1.7) *	12 (1.3)
Grade 12	Northeast	1994	54 (2.4)	33 (1.7)	12 (1.3)	1 (0.4)	46 (2.4)	13 (1.5)
		2001	55 (3.8)	31 (1.9)	11 (1.7)	2 (***)	45 (3.8)	13 (3.2)
	Southeast	1994	63 (1.9)	29 (1.6)	8 (1.3)	# (0.3)	37 (1.9)	8 (1.4)
		2001	61 (2.3)	29 (1.4)	9 (1.2)	1 (0.3)	39 (2.3)	10 (1.3)
	Central	1994	55 (2.0)	34 (1.3)	10 (1.1)	1 (0.5)	45 (2.0)	11 (1.2)
		2001	54 (2.2)	35 (2.0)	10 (1.2)	1 (0.4)	46 (2.2)	11 (1.3)
	West	1994	57 (2.3)	33 (2.0)	9 (1.2)	1 (0.2)	43 (2.3)	10 (1.2)
		2001	58 (2.2)	31 (1.5)	10 (1.2)	1 (0.4)	42 (2.2)	11 (1.5)

Standard errors of the estimated percentages appear in parentheses.
* Significantly different from 1994.
# Percentage is between 0.0 and 0.5.
(***) Standard error estimates cannot be accurately determined.
NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 History Assessments.

Table B.12: Data for Figure 3.9 National Scale Score Results by Parents' Education

Percentage of students and average U.S. history scale scores by parents' highest level of education, grades 8 and 12: 1994 and 2001

		Less than high school	Graduated high school	Some education after high school	Graduated college	Unknown
Grade 8	1994	7 (0.4) 241 (1.3)	23 (0.8) 251 (0.8)	19 (0.5) 264 (0.8)	42 (1.0) 270 (0.8)	9 (0.4) 238 (1.4)
	2001	7 (0.5) 243 (2.3)	18 (0.7) 253 (1.1)	19 (0.5) 265 (1.0)	47 (1.2) 275 (0.8) *	9 (0.4) 244 (1.5) *
Grade 12	1994	7 (0.4) 263 (1.4)	20 (0.7) 276 (1.1)	25 (0.7) 287 (1.2)	45 (1.0) 296 (0.9)	3 (0.2) 256 (2.7)
	2001	7 (0.4) 269 (1.5) *	19 (0.6) 274 (1.0)	25 (0.7) 286 (0.8)	46 (1.2) 298 (1.3)	3 (0.2) 262 (2.4)

The percentage of students is listed first with the corresponding average scale score presented below.
Standard errors of the estimated percentages and scale scores appear in parentheses.
* Significantly different from 1994.
NOTE: Percentage may not add to 100 due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

Table B.13: Data for Figure 3.10 National Achievement-Level Results by Parents' Education

Percentage of students within and at or above U.S. history achievement levels by parents' highest level of education, grades 8 and 12: 1994 and 2001

		Below *Basic*	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
Grade 8							
Less than high school	1994	63 (2.2)	34 (2.0)	3 (0.7)	0 (***)	37 (2.2)	3 (0.7)
	2001	59 (3.3)	38 (3.2)	3 (1.7)	# (***)	41 (3.3)	3 (1.8)
Graduated high school	1994	50 (1.4)	44 (1.5)	6 (0.9)	# (0.1)	50 (1.4)	7 (0.9)
	2001	48 (1.7)	44 (1.6)	7 (1.0)	# (***)	52 (1.7)	7 (1.0)
Some education after high school	1994	32 (1.3)	54 (1.5)	13 (1.2)	# (0.2)	68 (1.3)	14 (1.1)
	2001	30 (1.3)	56 (1.6)	14 (1.3)	1 (0.3)	70 (1.3)	14 (1.3)
Graduated college	1994	26 (1.2)	53 (1.2)	20 (1.1)	1 (0.3)	74 (1.2)	22 (1.1)
	2001	22 (1.0)	52 (0.9)	24 (1.1)	3 (0.5) *	78 (1.0)	27 (1.1) *
Unknown	1994	64 (2.2)	33 (1.8)	3 (1.1)	# (***)	36 (2.2)	3 (1.1)
	2001	59 (2.6)	37 (3.0)	4 (1.1)	# (***)	41 (2.6)	4 (1.2)
Grade 12							
Less than high school	1994	85 (1.9)	14 (1.8)	1 (0.7)	# (***)	15 (1.9)	1 (0.6)
	2001	80 (2.1)	18 (1.9)	2 (0.7)	# (***)	20 (2.1)	2 (0.7)
Graduated high school	1994	71 (1.5)	24 (1.3)	4 (0.8)	# (***)	29 (1.5)	4 (0.8)
	2001	74 (1.3)	22 (1.3)	4 (0.8)	# (0.1)	26 (1.3)	4 (0.8)
S...e education after high school	1994	58 (1.9)	34 (2.0)	7 (1.1)	1 (0.3)	42 (1.9)	8 (1.1)
	2001	61 (1.3)	31 (1.2)	8 (0.7)	1 (0.2)	39 (1.3)	8 (0.7)
Graduated college	1994	44 (1.3)	39 (1.1)	15 (1.0)	1 (0.4)	56 (1.3)	17 (1.0)
	2001	42 (1.5)	40 (1.3)	16 (1.0)	2 (0.8)	58 (1.5)	18 (1.5)
Unknown	1994	88 (3.5)	11 (3.2)	1 (***)	# (***)	12 (3.5)	1 (***)
	2001	83 (3.3)	14 (2.9)	3 (1.5)	# (***)	17 (3.3)	3 (1.5)

Standard errors of the estimated percentages appear in parentheses.
* Significantly different from 1994.
# Percentage is between 0.0 and 0.5.
(***) Standard error estimates cannot be accurately determined.
NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

BEST COPY AVAILABLE

Table B.14: Data for Figure 3.11 National Scale Score Results by Type of School

Percentage of students and average U.S. history scale scores by type of school, grades 4, 8, and 12: 1994 and 2001

		Public	Nonpublic	Nonpublic: Catholic	Nonpublic: Other
Grade 4	1994	90 (0.8) 203 (1.2)	10 (0.8) 222 (1.9)	6 (0.7) 221 (2.5)	4 (0.5) 224 (3.1)
	2001	88 (1.1) 207 (1.2) *	12 (1.1) 226 (1.9)	6 (0.7) 229 (2.5) *	5 (0.9) 223 (2.8)
Grade 8	1994	90 (0.9) 257 (0.7)	10 (0.9) 278 (1.1)	6 (0.6) 279 (1.5)	4 (0.6) 277 (2.1)
	2001	90 (0.9) 260 (0.8) *	10 (0.9) 279 (2.4)	5 (0.6) 280 (1.8)	5 (0.7) 278 (4.7)
Grade 12	1994	89 (1.1) 284 (0.8)	11 (1.1) 299 (1.3)	6 (0.9) 298 (2.2)	5 (0.6) 299 (2.2)
	2001	93 (0.8) 286 (1.1)	7 (0.8) 298 (2.0)	4 (0.6) 302 (2.4)	3 (0.5) 293 (3.1)

The percentage of students is listed first with the corresponding average scale score presented below.
Standard errors of the estimated percentages and scale scores appear in parentheses.
* Significantly different from 1994.
NOTE: Percentages may not add to 100 due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

Table B.15: Data for Figure 3.12 National Achievement-Level Results by Type of School

Percentage of students within and at or above U.S. history achievement levels by type of school, grades 4, 8, and 12: 1994 and 2001

			Below *Basic*	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
Grade 4	Public	1994	38 (1.2)	46 (1.0)	14 (1.0)	2 (0.3)	62 (1.2)	16 (1.1)
		2001	35 (1.3)	48 (1.3)	15 (1.0)	2 (0.5)	65 (1.3)	17 (1.1)
	Nonpublic	1994	18 (2.1)	55 (1.7)	23 (1.8)	3 (0.7)	82 (2.1)	26 (1.9)
		2001	15 (2.1)	55 (2.8)	26 (2.4)	4 (1.2)	85 (2.1)	30 (2.7)
	Nonpublic: Catholic	1994	19 (2.6)	56 (2.0)	22 (2.2)	2 (0.7)	81 (2.6)	24 (2.3)
		2001	14 (2.1)	51 (2.7)	29 (2.2)	5 (1.7)	86 (2.1)	35 (2.9) *
	Nonpublic: Other	1994	17 (4.0)	54 (3.6)	25 (3.5)	5 (1.4)	83 (4.0)	29 (3.9)
		2001	16 (4.0)	59 (6.2)	22 (4.4)	3 (1.3)	84 (4.0)	25 (4.9)
Grade 8	Public	1994	41 (1.0)	47 (0.8)	11 (0.6)	1 (0.1)	59 (1.0)	12 (0.6)
		2001	38 (1.0)	47 (1.1)	14 (0.8)	1 (0.3)	62 (1.0)	15 (0.8) *
	Nonpublic	1994	16 (1.2)	57 (1.8)	25 (1.7)	2 (0.6)	84 (1.2)	28 (1.8)
		2001	16 (3.2)	53 (2.5)	28 (2.1)	3 (0.9)	84 (3.2)	31 (2.6)
	Nonpublic: Catholic	1994	15 (1.7)	57 (2.1)	26 (2.1)	2 (0.8)	85 (1.7)	29 (2.3)
		2001	15 (2.0)	55 (1.7)	28 (2.1)	3 (0.8)	85 (2.0)	31 (2.4)
	Nonpublic: Other	1994	17 (2.8)	57 (3.1)	24 (3.0)	2 (0.8)	83 (2.8)	26 (3.4)
		2001	18 (6.2)	51 (4.7)	27 (3.9)	4 (1.5)	82 (6.2)	31 (4.9)
Grade 12	Public	1994	59 (1.2)	31 (0.9)	9 (0.7)	1 (0.2)	41 (1.2)	10 (0.7)
		2001	58 (1.3)	31 (0.9)	9 (0.7)	1 (0.4)	42 (1.3)	11 (1.0)
	Nonpublic	1994	41 (2.2)	41 (1.7)	17 (1.3)	1 (0.4)	59 (2.2)	18 (1.3)
		2001	41 (2.8)	42 (2.0)	15 (2.0)	2 (0.4)	59 (2.8)	17 (2.1)
	Nonpublic: Catholic	1994	43 (3.8)	40 (2.7)	17 (1.9)	1 (0.4)	57 (3.8)	18 (2.0)
		2001	36 (3.3)	44 (2.2)	18 (2.4)	2 (0.7)	64 (3.3)	20 (2.7)
	Nonpublic: Other	1994	39 (2.9)	42 (2.1)	17 (2.5)	2 (0.8)	61 (2.9)	19 (2.7)
		2001	47 (4.7)	39 (3.7)	12 (2.8)	1 (0.6)	53 (4.7)	13 (2.9)

Standard errors of the estimated percentages appear in parentheses.

* Significantly different from 1994.

NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.

SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

BEST COPY AVAILABLE

Table B.16: Data for Table 3.1 National Scale Score Results by Type of Location

Percentage of students and average U.S. history scale scores by type of location, grades 4, 8, and 12: 2001

	Central city	Urban fringe/large town	Rural/small town
Grade 4	27 (1.6) 199 (2.4)	45 (2.9) 211 (2.1)	28 (2.8) 215 (1.7)
Grade 8	27 (2.0) 257 (1.8)	45 (2.7) 265 (1.4)	28 (2.5) 263 (1.4)
Grade 12	25 (1.9) 283 (1.7)	40 (2.8) 292 (2.0)	35 (2.6) 284 (1.0)

The percentage of students is listed first with the corresponding average scale score presented below.
Standard errors of the estimated percentages and scale scores appear in parentheses.
NOTE: Percentages may not add to 100 due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Table B.17: Data for Figure 3.13 National Achievement-Level Results by Type of Location

Percentage of students within and at or above U.S. history achievement levels by type of location, grades 4, 8, and 12: 2001

		Below *Basic*	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
Grade 4	Central city	45 (2.5)	41 (2.0)	13 (1.8)	2 (0.5)	55 (2.5)	15 (2.0)
	Urban fringe/large town	31 (2.1)	49 (1.8)	17 (1.6)	3 (0.9)	69 (2.1)	20 (1.8)
	Rural/small town	23 (2.3)	58 (1.9)	17 (1.8)	2 (0.8)	77 (2.3)	19 (2.2)
Grade 8	Central city	43 (2.1)	42 (1.4)	13 (1.2)	2 (0.4)	57 (2.1)	15 (1.3)
	Urban fringe/large town	33 (1.8)	49 (1.2)	17 (1.0)	1 (0.4)	67 (1.8)	18 (1.1)
	Rural/small town	34 (1.9)	51 (1.8)	13 (1.9)	2 (0.5)	66 (1.9)	15 (1.8)
Grade 12	Central city	61 (2.2)	29 (1.7)	9 (0.8)	1 (0.2)	39 (2.2)	10 (0.9)
	Urban fringe/large town	52 (2.3)	33 (1.4)	13 (1.4)	2 (0.9)	48 (2.3)	15 (2.0)
	Rural/small town	60 (1.8)	32 (1.6)	7 (0.7)	# (0.2)	40 (1.8)	7 (0.7)

Standard errors of the estimated percentages appear in parentheses.
Percentage is between 0.0 and 0.5.
NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Table B.18: Data for Table 3.2 National Scale Score Results by Free/Reduced-Price School Lunch Program Eligibility

Percentage of students and average U.S. history scale scores by student eligibility for Free/Reduced-Price School Lunch program, grades 4, 8, and 12: 2001

	Eligible	Not eligible	Info not available
Grade 4	33 (1.4) 189 (1.6)	48 (2.1) 220 (1.4)	19 (2.4) 217 (2.8)
Grade 8	25 (1.1) 245 (1.2)	54 (2.1) 269 (0.9)	21 (2.2) 268 (2.0)
Grade 12	16 (0.9) 271 (1.3)	64 (2.2) 289 (1.2)	21 (2.5) 295 (2.0)

The percentage of students is listed first with the corresponding average scale score presented below.
Standard errors of the estimated percentages and scale scores appear in parentheses.
NOTE: Percentages may not add to 100 due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Table B.19: Data for Figure 3.14 National Achievement-Level Results by Free/Reduced-Price School Lunch Program Eligibility

Percentage of students within and at or above U.S. history achievement levels by student eligibility for the Free/Reduced-Price School Lunch program, grades 4, 8, and 12: 2001

		Below *Basic*	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
Grade 4	Eligible	53 (1.5)	40 (1.5)	6 (0.7)	1 (0.2)	47 (1.7)	6 (0.8)
	Not eligible	21 (1.7)	55 (2.0)	21 (1.4)	3 (0.8)	79 (1.7)	25 (1.6)
	Info not available	25 (2.8)	51 (2.5)	21 (2.4)	3 (1.1)	75 (2.8)	24 (2.9)
Grade 8	Eligible	59 (1.4)	35 (1.4)	5 (0.7)	# (0.2)	41 (1.4)	6 (0.7)
	Not eligible	27 (1.2)	53 (1.3)	18 (1.2)	2 (0.3)	73 (1.2)	20 (1.2)
	Info not available	30 (2.4)	48 (1.5)	19 (1.7)	3 (0.6)	70 (2.4)	22 (2.1)
Grade 12	Eligible	77 (1.8)	19 (1.7)	3 (0.7)	# (***)	23 (1.8)	3 (0.7)
	Not eligible	55 (1.5)	34 (1.1)	10 (0.7)	1 (0.6)	45 (1.5)	11 (1.1)
	Info not available	47 (2.9)	36 (2.3)	16 (1.9)	2 (0.6)	53 (2.9)	17 (2.3)

Standard errors of the estimated percentages appear in parentheses.
Percentage is between 0.0 and 0.5.
(***) Standard error estimates cannot be accurately determined.
NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

BEST COPY AVAILABLE

Table B.20: Data for Table 4.1 Comparison of Two Sets of National Scale Score Results

National average U.S. history scale scores by type of results, grades 4, 8, and 12: 2001

	Accommodations not permitted	Accommodations permitted
Grade 4	209 (1.0)	208 (0.9)
Grade 8	262 (0.8)	260 (0.8) †
Grade 12	287 (1.0)	287 (0.9)

Standard errors of the estimated scale scores appear in parentheses.
† Significantly different from the sample where accommodations were not permitted.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Table B.21: Data for Table 4.2 Comparison of Two Sets of National Achievement-Level Results

Percentage of students within and at or above U.S. history achievement levels by type of results, grades 4, 8, and 12: 2001

	Below *Basic*	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
Grade 4						
Accommodations were not permitted	33 (1.1)	49 (1.1)	16 (0.9)	2 (0.5)	67 (1.1)	18 (1.0)
Accommodations were permitted	34 (1.2)	48 (1.1)	16 (0.9)	2 (0.3)	66 (1.2)	18 (0.9)
Grade 8						
Accommodations were not permitted	36 (0.9)	48 (0.9)	15 (0.8)	2 (0.3)	64 (0.9)	17 (0.8)
Accommodations were permitted	38 (1.0) †	46 (0.9) †	14 (0.7)	1 (0.2)	62 (1.0) †	16 (0.7)
Grade 12						
Accommodations were not permitted	57 (1.2)	32 (0.9)	10 (0.6)	1 (0.4)	43 (1.2)	11 (0.9)
Accommodations were permitted	57 (1.2)	32 (0.9)	10 (0.7)	1 (0.3)	43 (1.2)	11 (0.9)

Standard errors of the estimated percentages appear in parentheses.
† Significantly different from the sample where accommodations were not permitted.
NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

BEST COPY AVAILABLE

Table B.22: Data for Table 4.3 Comparison of Two Sets of National Scale Score Results by Gender

National average U.S. history scale scores by gender and type of results, grades 4, 8, and 12: 2001

	Male	Female
Grade 4		
Accommodations were not permitted	209 (1.1)	209 (1.2)
Accommodations were permitted	207 (1.1)	209 (1.2)
Grade 8		
Accommodations were not permitted	264 (0.9)	261 (0.9)
Accommodations were permitted	261 (0.9) †	260 (0.9) †
Grade 12		
Accommodations were not permitted	288 (1.3)	286 (0.9)
Accommodations were permitted	288 (1.1)	286 (0.9)

Standard errors of the estimated scale scores appear in parentheses.
† Significantly different from the sample where accommodations were not permitted.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Table B.23: Data for Table 4.4 Comparison of Two Sets of National Achievement-Level Results by Gender

Percentage of students within and at or above U.S. history achievement levels by gender and type of results, grades 4, 8, and 12: 2001

	Below *Basic*	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
Grade 4						
Male						
Accommodations were not permitted	34 (1.3)	47 (1.3)	17 (1.1)	2 (0.7)	66 (1.3)	19 (1.2)
Accommodations were permitted	35 (1.3)	46 (1.3)	16 (1.3)	2 (0.5)	65 (1.3)	19 (1.3)
Female						
Accommodations were not permitted	32 (1.4)	51 (1.6)	15 (1.1)	2 (0.4)	68 (1.4)	17 (1.1)
Accommodations were permitted	33 (1.4)	50 (1.5)	15 (1.3)	2 (0.4)	67 (1.4)	17 (1.2)
Grade 8						
Male						
Accommodations were not permitted	35 (1.1)	47 (1.3)	17 (1.0)	2 (0.3)	65 (1.1)	18 (1.0)
Accommodations were permitted	38 (1.1) †	45 (1.0)	16 (0.9)	2 (0.3)	62 (1.1) †	17 (0.9)
Female						
Accommodations were not permitted	37 (1.2)	48 (1.0)	14 (0.8)	1 (0.4)	63 (1.2)	15 (0.8)
Accommodations were permitted	39 (1.2)	47 (1.2)	13 (0.9)	1 (0.2)	61 (1.2)	14 (0.9)
Grade 12						
Male						
Accommodations were not permitted	55 (1.6)	33 (1.1)	11 (0.8)	1 (0.5)	45 (1.6)	12 (1.1)
Accommodations were permitted	55 (1.6)	32 (1.3)	11 (0.9)	2 (0.4)	45 (1.6)	12 (1.1)
Female						
Accommodations were not permitted	59 (1.3)	31 (1.2)	9 (0.7)	1 (0.3)	41 (1.3)	10 (0.9)
Accommodations were permitted	60 (1.2)	31 (1.1)	9 (0.8)	1 (0.3)	40 (1.2)	10 (0.9)

Standard errors of the estimated percentages appear in parentheses.
† Significantly different from the sample where accommodations were not permitted.
NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

BEST COPY AVAILABLE

Table B.24: Data for Table 4.5 Comparison of Two Sets of National Scale Score Results by Race/Ethnicity

National average U.S. history scale scores by race/ethnicity and type of results, grades 4, 8, and 12: 2001

	White	Black	Hispanic	Asian/Pacific Islander	American Indian
Grade 4					
Accommodations were not permitted	220 (1.1)	188 (1.8)	196 (2.5)	213 (2.7)	197 (6.9)
Accommodations were permitted	218 (1.2)	186 (2.0)	187 (2.0)	214 (3.3)	197 (5.1)
Grade 8					
Accommodations were not permitted	271 (0.8)	243 (1.8)	243 (1.5)	267 (3.4)	249 (4.5)
Accommodations were permitted	269 (0.9) †	240 (1.8) †	240 (1.8)	265 (2.6)	248 (4.4)
Grade 12					
Accommodations were not permitted	292 (1.0)	269 (1.5)	274 (1.7)	295 (4.6)	277 (5.5) !
Accommodations were permitted	292 (0.9)	268 (1.4)	271 (1.8)	294 (5.5)	274 (5.5) !

Standard errors of the estimated scale scores appear in parentheses.
† Significantly different from the sample where accommodations were not permitted.
! The nature of the sample does not allow accurate determination of the variability of the statistic.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Table B.25: Data for Table 4.6 Comparison of Two Sets of National Achievement-Level Results by Race/Ethnicity

Percentage of students within and at or above U.S. history achievement levels by race/ethnicity and type of results, grades 4, 8, and 12: 2001

	Below *Basic*	At *Basic*	At *Proficient*	At *Advanced*	At or above *Basic*	At or above *Proficient*
Grade 4						
White						
Accommodations were not permitted	21 (1.3)	55 (1.4)	21 (1.4)	3 (0.7)	79 (1.3)	24 (1.4)
Accommodations were permitted	23 (1.5)	53 (1.3)	21 (1.2)	3 (0.5)	77 (1.5)	24 (1.3)
Black						
Accommodations were not permitted	56 (2.1)	38 (1.9)	5 (0.9)	# (0.3)	44 (2.1)	6 (1.0)
Accommodations were permitted	58 (2.3)	36 (2.0)	5 (1.0)	# (0.3)	42 (2.3)	5 (0.9)
Hispanic						
Accommodations were not permitted	58 (3.0)	35 (2.6)	6 (1.1)	1 (0.3)	42 (3.0)	7 (1.1)
Accommodations were permitted	58 (2.0)	36 (1.8)	6 (0.7)	# (0.3)	42 (2.0)	6 (0.8)
Asian/Pacific Islander						
Accommodations were not permitted	29 (3.8)	53 (4.6)	16 (2.7)	3 (1.9)	71 (3.8)	19 (3.2)
Accommodations were permitted	26 (4.0)	54 (4.2)	17 (2.6)	3 (***)	74 (4.0)	20 (3.2)
American Indian						
Accommodations were not permitted	47 (6.4)	41 (6.0)	8 (3.0)	4 (***)	53 (6.4)	12 (4.6)
Accommodations were permitted	44 (5.4)	44 (4.7)	9 (3.0)	3 (***)	56 (5.4)	12 (4.2)
Grade 8						
White						
Accommodations were not permitted	25 (1.0)	53 (1.1)	19 (1.1)	2 (0.4)	75 (1.0)	21 (1.1)
Accommodations were permitted	27 (1.1)	52 (0.9)	19 (0.9)	2 (0.3)	73 (1.1)	20 (1.0)
Black						
Accommodations were not permitted	62 (2.4)	34 (2.1)	4 (0.8)	# (***)	38 (2.4)	4 (0.8)
Accommodations were permitted	65 (2.1)	31 (1.7)	4 (0.6)	# (***)	35 (2.1)	4 (0.7)
Hispanic						
Accommodations were not permitted	60 (1.7)	34 (1.5)	5 (0.7)	# (0.2)	40 (1.7)	5 (0.7)
Accommodations were permitted	63 (2.2)	32 (1.9)	4 (0.7)	# (0.2)	37 (2.2)	4 (0.6)
Asian/Pacific Islander						
Accommodations were not permitted	32 (3.8)	48 (3.0)	18 (3.2)	2 (0.8)	68 (3.8)	20 (3.6)
Accommodations were permitted	34 (3.1)	47 (3.1)	17 (3.4)	2 (1.0)	66 (3.1)	19 (3.2)
American Indian						
Accommodations were not permitted	50 (7.1)	42 (6.7)	7 (3.2)	1 (***)	50 (7.1)	8 (3.5)
Accommodations were permitted	54 (6.0)	38 (5.4)	7 (2.2)	1 (***)	46 (6.0)	8 (2.4)
Grade 12						
White						
Accommodations were not permitted	51 (1.4)	36 (1.1)	12 (0.8)	1 (0.4)	49 (1.4)	13 (1.0)
Accommodations were permitted	51 (1.4)	36 (1.0)	12 (0.9)	1 (0.3)	49 (1.4)	13 (1.0)
Black						
Accommodations were not permitted	80 (1.5)	18 (1.4)	3 (0.6)	# (***)	20 (1.5)	3 (0.6)
Accommodations were permitted	80 (1.6)	17 (1.4)	3 (0.5)	# (***)	20 (1.6)	3 (0.5)
Hispanic						
Accommodations were not permitted	74 (2.4)	21 (1.8)	5 (1.0)	# (***)	26 (2.4)	5 (1.1)
Accommodations were permitted	74 (2.1)	21 (1.5)	5 (0.9)	# (0.2)	26 (2.1)	5 (1.0)
Asian/Pacific Islander						
Accommodations were not permitted	47 (5.1)	31 (2.4)	17 (4.1)	5 (2.3)	53 (5.1)	21 (6.0)
Accommodations were permitted	48 (5.8)	31 (2.4)	16 (4.8)	5 (2.7)	52 (5.8)	21 (6.6)
American Indian						
Accommodations were not permitted	66 (7.2) !	33 (7.4) !	1 (***)	0 (***)	34 (7.2) !	1 (***)
Accommodations were permitted	68 (7.8) !	31 (8.0) !	1 (***)	0 (***)	32 (7.8) !	1 (***)

Standard errors of the estimated percentages appear in parentheses.
Percentage is between 0.0 and 0.5.
(***) Standard error estimates cannot be accurately determined.
! The nature of the sample does not allow accurate determination of the variability of the statistic.
NOTE: Percentages within each U.S. history achievement-level range may not add to 100, or to the exact percentages at or above achievement levels, due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

BEST COPY AVAILABLE

Table B.26: Data for Table 5.1 Grade 4 Sample Question 1 Results (Multiple-Choice)

Overall percentage correct and percentages correct within each achievement-level range: 2001

Grade 4	Percentage correct within achievement-level intervals			
Overall percentage correct	Below *Basic* 194 and below*	*Basic* 195–242*	*Proficient* 243–275*	*Advanced* 276 and above*
93 (0.7)	84 (1.8)	96 (0.8)	99 (***)	*** (***)

Standard errors of the estimated percentages appear in parentheses.
*NAEP U.S. history composite scale range.
***(***) Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Table B.27: Data for Table 5.2 Grade 4 Sample Question 2 Results (Multiple-Choice)

Overall percentage correct and percentages correct within each achievement-level range: 2001

Grade 4	Percentage correct within achievement-level intervals			
Overall percentage correct	Below *Basic* 194 and below*	*Basic* 195–242*	*Proficient* 243–275*	*Advanced* 276 and above*
45 (1.3)	25 (2.0)	46 (2.2)	74 (3.3)	*** (***)

Standard errors of the estimated percentages appear in parentheses.
*NAEP U.S. history composite scale range.
***(***) Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Table B.28: Data for Table 5.3 Grade 4 Sample Question 3 Results (Multiple-Choice)

Overall percentage correct and percentages correct within each achievement-level range: 2001

Grade 4	Percentage correct within achievement-level intervals			
Overall percentage correct	Below *Basic* 194 and below*	*Basic* 195–242*	*Proficient* 243–275*	*Advanced* 276 and above*
57 (1.5)	34 (2.3)	59 (2.3)	86 (2.2)	*** (***)

Standard errors of the estimated percentages appear in parentheses.
*NAEP U.S. history composite scale range.
***(***) Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Table B.29. Data for Table 5.4 Grade 4 Sample Question 4 Results (Extended Constructed-Response)

Overall percentage "Essential" or better and percentages "Essential" or better within each achievement-level range: 2001

Grade 4	Percentage "Essential" or better within achievement-level intervals			
Overall percentage "Essential" or better	**Below *Basic* 194 and below***	***Basic* 195–242***	***Proficient* 243–275***	***Advanced* 276 and above***
42 (1.2)	13 (1.8)	48 (2.3)	76 (3.2)	*** (***)

Standard errors of the estimated percentages appear in parentheses.
*NAEP U.S. history composite scale range.
***(***) Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Table B.30. Data for Table 5.5 Grade 8 Sample Question 5 Results (Multiple-Choice)

Overall percentage correct and percentages correct within each achievement-level range: 2001

Grade 8	Percentage correct within achievement-level intervals			
Overall percentage correct	**Below *Basic* 251 and below***	***Basic* 252–293***	***Proficient* 294–326***	***Advanced* 327 and above***
52 (1.9)	34 (2.3)	56 (2.4)	79 (2.7)	*** (***)

Standard errors of the estimated percentages appear in parentheses.
*NAEP U.S. history composite scale range.
***(***) Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Table B.31. Data for Table 5.6 Grade 8 Sample Question 6 Results (Multiple-Choice)

Overall percentage correct and percentages correct within each achievement-level range: 2001

Grade 8	Percentage correct within achievement-level intervals			
Overall percentage correct	**Below *Basic* 251 and below***	***Basic* 252–293***	***Proficient* 294–326***	***Advanced* 327 and above***
39 (1.2)	29 (1.3)	39 (1.7)	62 (4.0)	*** (***)

Standard errors of the estimated percentages appear in parentheses.
*NAEP U.S. history composite scale range.
***(***) Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Table B.32: Data for Table 5.7 Grade 8 Sample Question 7 Results (Short Constructed-Response)

Overall percentage "Appropriate" and percentages "Appropriate" within each achievement-level range: 2001

Grade 8	Percentage "Appropriate" within achievement-level intervals			
Overall percentage "Appropriate"	Below *Basic* 251 and below*	*Basic* 252–293*	*Proficient* 294–326*	*Advanced* 327 and above*
30 (1.5)	9 (1.9)	34 (2.2)	64 (4.3)	*** (***)

Standard errors of the estimated percentages appear in parentheses.
*NAEP U.S. history composite scale range.
***(***) Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Table B.33: Data for Table 5.8 Grade 8 Sample Question 8 Results (Short Constructed-Response)

Overall percentage "Appropriate" and percentages "Appropriate" within each achievement-level range: 2001

Grade 8	Percentage "Appropriate" within achievement-level intervals			
Overall percentage "Appropriate"	Below *Basic* 251 and below*	*Basic* 252–293*	*Proficient* 294–326*	*Advanced* 327 and above*
33 (1.1)	11 (1.5)	38 (2.0)	64 (4.9)	*** (***)

Standard errors of the estimated percentages appear in parentheses.
*NAEP U.S. history composite scale range.
***(***) Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Table B.34: Data for Table 5.9 Grade 12 Sample Question 9 Results (Multiple-Choice)

Overall percentage correct and percentages correct within each achievement-level range: 2001

Grade 12	Percentage correct within achievement-level intervals			
Overall percentage correct	Below *Basic* 293 and below*	*Basic* 294–324*	*Proficient* 325–354*	*Advanced* 355 and above*
36 (1.3)	23 (1.6)	47 (2.3)	73 (4.8)	*** (***)

Standard errors of the estimated percentages appear in parentheses.
*NAEP U.S. history composite scale range.
***(***) Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

BEST COPY AVAILABLE

Table B.35: Data for Table 5.10 Grade 12 Sample Question 10 Results (Multiple-Choice)

Overall percentage correct and percentages correct within each achievement-level range: 2001

Grade 12	Percentage correct within achievement-level intervals			
Overall percentage correct	Below *Basic* 293 and below*	*Basic* 294–324*	*Proficient* 325–354*	*Advanced* 355 and above*
68 (1.3)	54 (1.7)	82 (2.1)	95 (1.8)	*** (***)

Standard errors of the estimated percentages appear in parentheses.
*NAEP U.S. history composite scale range.
***(***) Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Table B.36: Data for Table 5.11 Grade 12 Sample Question 11 Results (Extended Constructed-Response)

Overall percentage "Essential" or better and percentages "Essential" or better within each achievement-level range: 2001

Grade 12	Percentage "Essential" or better within achievement-level intervals			
Overall percentage "Essential" or better	Below *Basic* 293 and below*	*Basic* 294–324*	*Proficient* 325–354*	*Advanced* 355 and above*
21 (1.5)	4 (0.8)	30 (2.6)	74 (3.8)	*** (***)

Standard errors of the estimated percentages appear in parentheses.
*NAEP U.S. history composite scale range.
***(***) Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Table B.37: Data for Table 5.12 Grade 12 Sample Question 12 Results (Extended Constructed-Response)

Overall percentage "Essential" or better and percentages "Essential" or better within each achievement-level range: 2001

Grade 12	Percentage "Essential" or better within achievement-level intervals			
Overall percentage "Essential" or better	Below *Basic* 293 and below*	*Basic* 294–324*	*Proficient* 325–354*	*Advanced* 355 and above*
39 (1.7)	17 (1.6)	62 (3.1)	88 (3.5)	*** (***)

Standard errors of the estimated percentages appear in parentheses.
*NAEP U.S. history composite scale range.
***(***) Sample size insufficient to permit a reliable estimate (see appendix A).
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Table B.38: Data for Table 6.1 Time Spent on Fourth-Grade Social Studies

Percentage of students and average U.S. history scale scores by teachers' reports on the amount of time spent on social studies in a typical week at grade 4: 2001

	2001
Less than 30 minutes	2 (0.6) 191 (7.8) !
30 to 60 minutes	12 (1.1) 195 (2.5)
61 to 120 minutes	37 (2.1) 210 (1.6)
121 to 180 minutes	31 (2.5) 211 (2.2)
More than 180 minutes	19 (2.1) 218 (2.3)

The percentage of students is listed first with the corresponding average scale score presented below.
Standard errors of the estimated percentages and scale scores appear in parentheses.
! The nature of the sample does not allow accurate determination of the variability of the statistic.
NOTE: Percentages may not add to 100 due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

BEST COPY AVAILABLE

Table B.39: Data for Table 6.2 Use of State/Local Standards in Planning Instruction

Percentage of students and average U.S. history scale scores by teachers' reports on the use of state/local standards in planning instruction at grades 4 and 8: 2001

	2001
Grade 4	
Not at all	2 (0.4) 212 (5.3) !
Small extent	9 (1.1) 210 (2.5)
Moderate extent	23 (1.8) 206 (2.6)
Large extent	63 (2.3) 210 (1.4)
No standards for teaching social studies	3 (1.2) 224 (6.4) !
Grade 8	
Not at all	2 (0.6) 274 (4.4) !
Small extent	7 (1.6) 264 (3.6) !
Moderate extent	21 (2.3) 266 (1.9)
Large extent	69 (2.6) 262 (1.1)
No standards for teaching social studies	1 (0.4) 276 (6.4) !

The percentage of students is listed first with the corresponding average scale score presented below.
Standard errors of the estimated percentages and scale scores appear in parentheses.
! The nature of the sample does not allow accurate determination of the variability of the statistic.
NOTE: Percentages may not add to 100 due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Table B.40: Data for Table 6.3 Frequency of Fourth- and Eighth-Grade Classroom Activities

Percentage of students and average U.S. history scale scores by teachers' reports on frequency of classroom activities at grades 4 and 8: 1994 and 2001

	1994	2001
Grade 4		
Reading material from a textbook		
Almost every day	43 (2.6) 207 (2.2)	41 (2.2) 214 (1.4)
Once or twice a week	44 (2.6) 204 (1.3)	47 (2.2) 207 (1.8)
Once or twice a month	8 (1.4) 204 (4.6)	7 (1.3) 202 (3.4)
Never or hardly ever	5 (0.9) 204 (5.0)	5 (1.3) 209 (6.7) !
Reading extra material not in the regular textbook		
Almost every day	6 (1.2) 208 (4.4) !	9 (1.2) 210 (3.4)
Once or twice a week	33 (2.2) 205 (1.9)	44 (2.5) * 211 (1.5)
Once or twice a month	46 (2.5) 204 (1.9)	35 (2.4) * 208 (1.8)
Never or hardly ever	15 (1.7) 208 (3.2)	12 (1.5) 208 (2.8)
Using primary historical documents		
Almost every day	1 (0.4) *** (***)	1 (0.4) *** (***)
Once or twice a week	8 (1.2) 201 (4.5)	11 (1.3) 207 (2.6)
Once or twice a month	29 (1.9) 208 (2.1)	39 (2.2) * 212 (2.0)
Never or hardly ever	62 (2.0) 205 (1.3)	48 (1.9) * 208 (1.6)
Writing a report		
Almost every day	*** (***) *** (***)	1 (0.4) 194 (8.0) !
Once or twice a week	6 (1.2) 188 (4.2)	6 (1.0) 198 (3.4)
Once or twice a month	63 (2.3) 207 (1.4)	59 (2.1) 210 (1.4)
Never or hardly ever	31 (2.1) 205 (2.1)	34 (2.0) 210 (1.8)

See footnotes at end of table. ▷

Table B.40: Data for Table 6.3 Frequency of Fourth- and Eighth-Grade Classroom Activities (continued)

Percentage of students and average U.S. history scale scores by teachers' reports on frequency of classroom activities at grades 4 and 8: 1994 and 2001

	1994	2001
Grade 8		
Reading material from a textbook		
Almost every day	45 (3.6) 259 (1.3)	45 (2.6) 264 (1.3)
Once or twice a week	42 (3.2) 259 (1.5)	45 (2.5) 262 (1.1)
Once or twice a month	8 (1.4) 266 (2.1)	7 (1.4) 262 (3.1)
Never or hardly ever	5 (1.8) 265 (3.6) !	3 (0.6) 275 (4.0) !
Reading extra material not in the regular textbook		
Almost every day	3 (0.6) 254 (4.5)	7 (1.2) * 265 (3.1)
Once or twice a week	32 (2.8) 258 (1.5)	37 (2.5) 261 (1.4)
Once or twice a month	47 (2.4) 263 (1.2)	44 (2.8) 264 (1.2)
Never or hardly ever	17 (2.2) 258 (2.5)	13 (1.6) 264 (2.5)
Using primary historical documents		
Almost every day	2 (0.7) 268 (9.4) !	4 (0.6) 264 (4.0)
Once or twice a week	20 (2.2) 260 (2.0)	27 (2.3) 267 (1.6)
Once or twice a month	55 (2.3) 261 (1.1)	54 (3.1) 262 (1.0)
Never or hardly ever	23 (1.9) 258 (1.6)	16 (2.2) * 259 (1.9)
Writing a report		
Almost every day	# (0.2) *** (***)	1 (0.3) 255 (4.0) !
Once or twice a week	4 (0.9) 256 (4.0) !	7 (1.4) 266 (3.6) !
Once or twice a month	66 (2.5) 261 (0.9)	66 (2.8) 263 (0.9)
Never or hardly ever	30 (2.6) 259 (1.6)	27 (2.6) 263 (1.7)

The percentage of students is listed first with the corresponding average scale score presented below.
Standard errors of the estimated percentages and scale scores appear in parentheses.
! The nature of the sample does not allow accurate determination of the variability of the statistic.
*Significantly different from 1994.
***(***) Sample size is insufficient to permit a reliable estimate.
Percentage is between 0.0 and 0.5.
NOTE: Percentages may not add to 100 due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

Table B.41: Data for Table 6.4 Frequency of Twelfth-Grade Classroom Activities

Percentage of students and average U.S. history scale scores by students' reports on frequency of classroom activities at grade 12: 1994 and 2001

	1994	2001
Read material from a textbook		
About every day	40 (0.9) 289 (0.8)	44 (1.1) * 290 (1.0)
Once or twice a week	40 (0.9) 289 (0.9)	38 (0.9) 289 (1.2)
Once or twice a month	9 (0.5) 284 (1.9)	8 (0.4) 283 (2.4)
A few times a year	6 (0.4) 278 (1.8)	6 (0.3) 276 (1.9)
Never	6 (0.5) 268 (1.8)	5 (0.3) * 270 (2.6)
Read extra material not in the regular textbook		
About every day	9 (0.4) 288 (1.7)	10 (0.4) 290 (1.5)
Once or twice a week	30 (0.8) 289 (1.2)	31 (0.7) 291 (1.2)
Once or twice a month	24 (0.5) 291 (1.1)	25 (0.6) 290 (1.4)
A few times a year	18 (0.6) 288 (1.0)	17 (0.5) 289 (1.2)
Never	18 (0.7) 274 (1.3)	17 (0.5) 276 (1.1)

See footnotes at end of table. ▷

Table B.41: Data for Table 6.4 Frequency of Twelfth-Grade Classroom Activities (continued)

Percentage of students and average U.S. history scale scores by students' reports on frequency of classroom activities at grade 12: 1994 and 2001

	1994	2001
Use letters, diaries, or essays written by historical people		
About every day	4 (0.2) 280 (3.2)	5 (0.3) * 280 (2.1)
Once or twice a week	14 (0.5) 292 (1.5)	18 (0.6) * 290 (1.4)
Once or twice a month	24 (0.5) 291 (1.2)	26 (0.6) * 291 (1.1)
A few times a year	26 (0.5) 291 (0.9)	26 (0.5) 292 (1.2)
Never	32 (0.7) 279 (0.9)	25 (0.6) * 279 (0.9)
Write a report		
About every day	2 (0.2) 267 (3.6)	3 (0.2) * 271 (3.0)
Once or twice a week	9 (0.4) 279 (1.9)	14 (0.6) * 288 (1.4)
Once or twice a month	35 (0.7) 287 (0.9)	41 (0.8) * 290 (1.0)
A few times a year	40 (0.8) 293 (0.9)	34 (1.1) * 290 (1.1)
Never	14 (0.8) 278 (1.4)	8 (0.4) * 271 (1.7)

The percentage of students is listed first with the corresponding average scale score presented below.
Standard errors of the estimated percentages and scale scores appear in parentheses.
*Significantly different from 1994.
NOTE: Percentages may not add to 100 due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1994 and 2001 U.S. History Assessments.

BEST COPY AVAILABLE

Table B.42: Data for Table 6.5 Fourth-Grade Computer Use

Percentage of students and average U.S. history scale scores by students' reports on computer use at grade 4: 2001

	2001
Use computers at school for social studies	
Every day	3 (0.2) 167 (3.5)
Two or three times a week	5 (0.4) 186 (2.5)
Once a week	7 (0.5) 197 (3.9)
Once every few weeks	10 (0.5) 212 (2.4)
Never or hardly ever	74 (1.1) 214 (0.9)
Do research projects using a CD or the Internet	
Yes	46 (1.1) 211 (1.4)
No	54 (1.1) 208 (1.1)
Use computer to write reports	
Yes	52 (1.2) 209 (1.4)
No	48 (1.2) 210 (0.9)

The percentage of students is listed first with the corresponding average scale score presented below.
Standard errors of the estimated percentages and scale scores appear in parentheses.
NOTE: Percentages may not add to 100 due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

BEST COPY AVAILABLE

Table B.43: Data for Table 6.6 Eighth- and Twelfth-Grade Computer Use

Percentage of students and average U.S. history scale scores by students' reports on computer use at grades 8 and 12: 2001

	2001
Grade 8	
Use computers at school for social studies	
Every day	1 (0.2) 239 (3.7)
Two or three times a week	5 (0.4) 252 (2.9)
Once a week	9 (0.6) 261 (1.7)
Once every few weeks	21 (0.8) 268 (0.9)
Never or hardly ever	64 (1.2) 263 (0.9)
Do research projects using a CD or the Internet	
Not at all	26 (1.0) 253 (1.1)
Small extent	31 (0.7) 262 (0.8)
Moderate extent	26 (0.7) 267 (1.0)
Large extent	16 (0.7) 272 (1.1)
Write reports on the computer	
Not at all	18 (0.9) 253 (1.3)
Small extent	30 (0.7) 260 (0.9)
Moderate extent	29 (0.7) 266 (0.9)
Large extent	23 (0.8) 270 (1.1)

See footnotes at end of table ▷

Table B.43. Data for Table 8.6 Eighth- and Twelfth-Grade Computer Use (continued)

Percentage of students and average U.S. history scale scores by students' reports on computer use at grades 8 and 12: 2001

	2001
Grade 12	
Use computers at school for studying history	
Every day	2 (0.1) 265 (4.0)
Two or three times a week	6 (0.3) 277 (2.1)
Once a week	7 (0.5) 280 (1.6)
Once every few weeks	16 (0.7) 291 (1.5)
Never or hardly ever	42 (1.0) 289 (1.1)
Haven't studied history this year	27 (1.3) 289 (1.1)
Do research projects using a CD or the Internet	
Not at all	23 (0.7) 274 (1.0)
Small extent	33 (0.7) 286 (1.0)
Moderate extent	29 (0.6) 294 (1.2)
Large extent	15 (0.5) 300 (1.7)
Write reports on the computer	
Not at all	14 (0.5) 271 (1.1)
Small extent	27 (0.7) 281 (1.1)
Moderate extent	33 (0.6) 290 (1.0)
Large extent	26 (0.8) 300 (1.4)

The percentage of students is listed first with the corresponding average scale score presented below.
Standard errors of the estimated percentages and scale scores appear in parentheses.
NOTE: Percentages may not add to 100 due to rounding.
SOURCE: U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2001 U.S. History Assessment.

Appendix C
Members of the NAEP U.S. History Standing Committee

Carol Ruth Berkin
Graduate Center–CUNY
New York, NY

Paul Cohen
East Brunswick Public Schools
East Brunswick, NJ

Richard del Rio
Muirlands Middle School
La Jolla, CA

Anthony Edmonds
Ball State University–History Dept.
Muncie, IN

Sara Forte
Bowne-Munro Elementary School
East Brunswick, NJ

Dennis Gregg
George Washington High School
San Francisco, CA

Linda Levstik
University of Kentucky
Lexington, KY

Tedd Levy
Nathan Hale Middle School
Norwalk, CT

Louisa Moffitt
Marist School
Atlanta, GA

Lori Morton
Woodley Hills Elementary School
Alexandria, VA

Diane Palmer
Belmont, MA

John Patrick
Social Studies Development Center
Indiana University
Bloomington, IN

Joann Prewitt
Delaware State Dept. of Education
Dover, DE

Bruce VanSledright
College of Education
University of Maryland
College Park, MD

Allan Winkler
Department of History
Miami University
Oxford, OH

Acknowledgments

This report is the culmination of the efforts of many individuals who contributed their considerable knowledge, experience, and creativity to the NAEP 2001 U.S. history assessment. The assessment was a collaborative effort among staff from the National Center for Education Statistics (NCES), the National Assessment Governing Board (NAGB), Educational Testing Service (ETS), Westat, and NCS Pearson. Most importantly, NAEP is grateful to the students and school staff whose participation made the assessment possible.

The NAEP 2001 U.S. history assessment was funded through NCES, in the Office of Educational Research and Improvement of the U.S. Department of Education. The Deputy Commissioner of Education Statistics, Gary W. Phillips, and the NCES staff—Janis Brown, Peggy Carr, Arnold Goldstein, Steven Gorman, Carol Johnson, and Andrew Kolstad—worked closely and collegially with the authors to produce this report.

The NAEP project at ETS is directed by Stephen Lazer and John Mazzeo, with assistance from John Barone. Sampling and data collection activities were conducted by Westat under the direction of Rene Slobasky, Nancy Caldwell, Keith Rust, and Dianne Walsh. Printing, distribution, scoring, and processing activities were conducted by NCS Pearson under the direction of Brad Thayer, William Buckles, Mathilde Kennel, Linda Reynolds, and Connie Smith.

Test development activities were conducted by ETS under the direction of Madeline Goodman and Michael Lapp.

The complex statistical and psychometric activities necessary to report results for the NAEP 2001 U.S. history assessment were directed by Catherine Hombo and by Brenda Siok-Hoon Tay-Lim, and were advised by Nancy Allen. The analyses that produced results for the data for which student accommodations were permitted were supervised by Brenda Tay-Lim with assistance from Hui Deng.

The extensive data processing and computer programming activities underlying the statistical and psychometric analyses conducted at ETS are under the direction of David Freund, Edward Kulick, Bruce Kaplan, and Steven Isham. Data analyses presented in this report were managed by Tatyana Petrovicheva with assistance from Satwinder Thind, Vincent Weng, Norma Norris, Alfred Rogers, and Mei-Jang Lin. The complex database work for this assessment was managed by Katherine Pashley with assistance from Gerry Kokolis.

The design and production of this report was overseen by Loretta Casalaina. Joseph Kolodey and Rick Hasney contributed invaluable design and production expertise to the effort. Wendy Grigg coordinated the documentation and data checking procedures with assistance from Janice Goodis, Andrea Bergen, and Alice Kass. Shari Santapau coordinated the editorial and proofreading procedures with assistance from Valerie Mukuna and Julie McGuire. The Web version of this report was coordinated by Pat O'Reilly with assistance from Rick Hasney.

Many thanks are due to the numerous reviewers, both internal and external to NCES and ETS. The comments and critical feedback of the following reviewers are reflected in the final version of this report: Janis Brown, Jay Campbell, James Carlson, Mary Crovo, Young Chun, Douglas Cochrane, Mary Crovo, Arnold Goldstein, Steven Gorman, Ellen Harkavy, Dana Kelly, Andrew Kolstad, Carolyn Lee, Ted Levy, John Patrick, Marianne Perie, Michael Petrilli, Patrick Rooney, John Sietsema, Rod Schwartz, Allen Vanneman, and Emily Wurtz.

United States
Department of Education
ED Pubs
8242-B Sandy Court
Jessup, MD 20794-1398

Postage and Fees Paid
U.S. Department of Education
Permit No. G-17

Parcel

Official Business Only
Penalty for Private Use, $300